UNIVERSITY OF PADUA

PH.D. SCHOOL IN BRAIN, MIND AND COMPUTER SCIENCE
Curriculum in Neuroscience, Technology and Society
XXXV Cycle

Cross-domain and cross-sensory effects of rhythm on language development

---

Candidate:
Sofia RUSSO
University of Padua

Supervisor:
Prof. Eloisa Valenza
University of Padua

Co-Supervisor:
Prof. Antonio Rodà
University of Padua

“L’emissione vocale e la scrittura seguissero dunque senza interruzione questo suo [del pensiero] nascere e rinascere.”

"Spazi metrici", Amelia Rosselli.

Abstract

Rhythmic abilities are deeply rooted in the human brain. They originated from our need for social interaction and bonding and have evolved hand-to-hand with communication and linguistic abilities across ages. In the ontogenetic evolution, first experiences with rhythm occur already in the womb, with different rhythmic sources being available to the human fetus. Among sensory modalities, vestibular, tactile, and somatosensory perception (Provasi et al., 2014, VTS) seems to play a crucial role in early rhythm processing, with neural coupling dynamics allowing for VTS and sensorimotor influences on auditory encoding (Trainor et al., 2009; Phillips-Silver and Trainor, 2005; Tichko et al., 2021). However, a restricted corpus of studies specifically focused on VTS and sensorimotor abilities in rhythm and language development. Therefore, the present work aimed at investigating the role of the body through VTS and sensorimotor rhythmic abilities in language development. Specifically, VTS rhythmic abilities will be firstly assessed through a custom made, vibrotactile tool for music perception specifically designed for infants and toddlers based on the recent advances in the field of human-computer interaction, as illustrated in Study 1 (Section 4.1). In Study 2 (Section 4.2), early linguistic abilities will be investigated in the same cohort of infants, specifically testing the role of phonological and prosodic features of speech. Then, the link between rhythmic and linguistic abilities will be explored between the two tasks, to investigate whether the former might be informative about the latter. In Study 3 (Section 4.3), cross-domain (i.e, from music to language) benefits resulting from rhythmic exposure across sensory modalities (i.e., from VTS to auditory signals) will be tested through a rhythmic priming paradigm. Therefore, Chapter 4 will investigate VTS rhythmic abilities and their link as well as their effect on the emerging linguistic abilities of infants and toddlers. In the following Chapter, the effect of sensorimotor rhythmic abilities on complex linguistic behaviors including reading and writing skills will be examined in young adults (Chapter 5).

Specifically, the effect of sensorimotor synchronization to an external beat will be evaluated while administering two screening tasks for developmental dyslexia: lexical decision and dictation under phonological suppression in Study 4 (Section 5.1) and Study 5 (Section 5.2). Therefore, Chapter 5 will provide insights into the role of sensorimotor rhythmic abilities in literacy skills with possible implications for early screening and training practices for neurodevelopmental disorders. Together, Chapter 4 and Chapter 5 together will bring evidence on the role of VTS experiences and sensorimotor abilities across language development. In conclusion, Chapter 6 will provide a general interpretation and discussion of the present findings with original insights on the cross-domain and cross-sensory effects of rhythm in language development.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Rhythmic abilities are building blocks for communication, social interaction, and language. Accordingly, humans are known to synchronize and move to rhythms, engage in social rhythmic activities (such as dance and choring), and enjoy the rhythmic structure of musical signals across cultures. Regarding communication, human languages are built on rhythmic properties similar to those of musical signals. Since birth, humans are sensitive to and rely on such rhythmic and temporal regularities to develop the capacity to express and communicate with others. In order to fully understand the cross-domain link between rhythm, communication, and language, Chapter 2 will initially report a definition of rhythm describing its sub-components and its structure in music and speech signals. Then, the evolutionary roots of rhythmic abilities will be briefly illustrated to highlight the role of rhythm in the phylogenetic development of communication and language. In the second section of Chapter 2, the neural and cognitive mechanisms underlying rhythm processing in humans will be illustrated, according to the nowadays models of Cognitive Neuroscience. Lastly, the ontogenetic roots of rhythm perception in the human womb will be described in the third section, exploring the link between prenatal experiences and sensory processing of rhythm throughout life. Specifically, the cross-sensory nature of early rhythmic experience will be discussed. Accordingly, auditory input in rhythm perception has been extensively studied; however, hearing is not the unique sensory modality involved in this process. Specifically, first experiences with rhythm, occurring already in the womb, encompass multiple sources of rhythmic stimulation, including the vestibular, tactile, and somatosensory input produced by maternal body's movements (Lecanuet and Schaal, 2002; Provasi et al., 2014). Consequently, vestibular and somatosensory inputs emerge as core aspects of rhythm perception after birth and during development (Trainor et al., 2009; Phillips-Silver and Trainor, 2005; Tichko et al., 2021).

Therefore, for their centrality in rhythm and language development, vestibular and somatosensory aspects of rhythm perception will be discussed in the main and last section of Chapter 2.

Interestingly, evidence suggests that early rhythmic abilities might be important for subsequent typical and atypical language development: this will be the main focus of Chapter 3. Specifically, linguistic rhythm is available to human fetuses since their first experience with language. At birth, signs of such early experiences can be found in their preference for familiar rhythms in language. Growing into infancy, early sensitivity to familiar rhythms guides perceptual and attentional processes toward the speech stream, at various levels of the prosodic hierarchy from metrical feet to phonological phrases. Therefore, linguistic rhythm is found to bootstrap early stages of language acquisition (Section 3.1). Interestingly, many parallels can be made between the hierarchical structure of language and music. Consequently, general rhythmic skills seem to be deeply related to language and music development (Section 3.2). This point, resulting from the last two decades of research, is in line with the claim made by the *processing rhythm in speech and music*framework (Fiveash et al., 2021, the PRISM framework, Chapter 2); that is, similar mechanisms underlie the processing of rhythmic signals mainly defined as precise auditory processing, oscillatory brain activity, and sensorimotor coupling. Despite the cognitive and neural dynamics underlying this cross-domain coupling is still unclear, further evidence in support of a domain-general set of rhythmic abilities serving the processing of multiple, complex rhythmic signals comes from the literature on atypical language development. Accordingly, atypical rhythmic skills have been found to be present across different neurodevelopmental disorders causing language- and motor-related difficulties (Goswami, 2011; Chang et al., 2016; Trainor et al., 2018; Carrer, 2015). Unifying these findings, the *atypical rhythm risk hypothesis* was formulated, stating that - given the crucial role of rhythmic abilities in language development - atypical rhythm could be a risk factor for developmental speech/language difficulties (Ladányi et al., 2020, ARRH). This promising hypothesis and the evidence that contributed to framing in will be discussed in Chapter 3 (Section 3.2).

Consistently with the ARRH, several findings demonstrated that short- and long-term music exposure can enhance cognitive and language performances (Miendlarzewska and Trost, 2014; Russo and Valenza, 2021). Specifically, the case of music education captured the interest of many researchers in the field of music and language cognition (Musacchia et al., 2007; Schön and Tillmann, 2015; Brown, 2012). Moreover, this line of research found increased basic auditory, cognitive, and high-functioning linguistic skills in musicians compared to individuals not receiving any professional musical education. This result led to investigating the mechanisms to transfer the effects of music on linguistic skills. The effects of exposure to musical rhythms on language processing have therefore been investigated, showing significant evidence for cross-domain transfers at multiple levels (Cason and Schön, 2012; Fiveash et al., 2021; Bedoin et al., 2016, e.g., phonological awareness, reading abilities, and grammar processing). Surprisingly, the effect of exposure to musical rhythms was also found to occur across sensory modalities (Fotidzis et al., 2018; Liu et al., 2012; Gould et al., 2018).

Therefore, - across Chapters - it becomes clear the main claim of this work; that is, rhythmic abilities and VTS/sensorimotor experiences might be foundational to language, from the earliest stages of acquisition and throughout the following developmental steps. The general idea that rhythm and movement are deeply connected is not new. Despite being relatively underexplored in developmental sciences (Bremner and Spence, 2017, Section 2.3), it is in fact common across anthropological theories and pedagogical techniques, offering the interesting insights illustrated in Section 3.3.3. In conclusion, the literature review constituting the Theoretical Section of this work aims to highlight how, despite its central role in rhythm perception, VTS experiences and sensorimotor abilities are still less investigated compared to other sensory modalities; yet, the strong prompt of the body experience in perceiving rhythm has been included in early music pedagogy with even more zeal than nowadays techniques. This is surprising considering that VTS experiences and sensorimotor abilities are natural channels for infants and children in investigating the world, given that they are highly familiar since the earliest experiences in the womb. Furthermore, VTS experiences and sensorimotor abilities can be crucial for infants and children with sensory deprivations (e.g., early blind or deaf individuals) as well as for people struggling with sensory processing (i.e., individuals with ASD or ADHD).

For these and many other reasons illustrated throughout this dissertation, VTS experiences and sensorimotor abilities in processing rhythm are worthy of further investigation. This is the prompt motivating the present work. Therefore, the Experimental Section aimed at investigating the role of the body through VTS and sensorimotor rhythmic abilities in language development. Specifically, VTS rhythmic abilities will be firstly assessed through the *Infant vibrotactile system for music perception*, specifically designed for infants and toddlers based on the recent advances in the field of human-computer interaction, as illustrated in Study 1 (section 4.1). In Study 2, section (4.2), early linguistic abilities will be investigated in the same cohort of infants, specifically testing the role of phonological and prosodic features of speech processing. Then, the link between rhythmic and linguistic abilities will be explored between the two tasks, to investigate whether the former might be informative about the latter. In Study 3 (section 4.3), cross-domain (i.e, from music to language) benefits resulting from rhythmic exposure across sensory modalities (i.e., from VTS to auditory signals) will be tested through a rhythmic priming paradigm. Therefore, Chapter 4 will investigate VTS rhythmic abilities and their link as well as their effect on the emerging linguistic abilities of infants and toddlers. In the following Chapter, the effect of sensorimotor rhythmic abilities on complex linguistic behaviors including reading and writing skills will be examined in young adults (Chapter 5). Specifically, the effect of sensorimotor synchronization to an external beat will be evaluated while administering two screening tasks for developmental dyslexia (lexical decision and dictation under phonological suppression; Study 4 in section 5.1 and Study 5 in section 5.2). Therefore, Chapter 5 will provide insights into the role of sensorimotor rhythmic abilities in literacy skills, with possible implications for early screening and training practices for neurodevelopmental disorders. In conclusion, Chapter 4 and Chapter 5 together will bring evidence on the role of VTS experiences and sensorimotor abilities across language development.

Part I

THEORETICAL FRAMEWORK

Chapter 2

Rhythm in cognition, communication, and language

Rhythmic abilities are building blocks for communication and social interaction in many different species, including humans. Accordingly, across cultures, humans are known to synchronize and move to rhythms, engage in social rhythmic activities (such as dance and choring), and enjoy the rhythmic structure of musical signals. Regarding communication, human languages are built on rhythmic properties similar to those of musical signals. Since birth, humans are sensitive to and rely on such rhythmic and temporal regularities to develop the capacity to express and communicate with others. In order to fully understand the deep link between rhythm, communication, and language, in the first section of this Chapter I will initially report a definition of rhythm describing its sub-components and its structure in music and speech signals. Then, the evolutionary roots of rhythmic abilities will be briefly illustrated to highlight the role of rhythm in the phylogenetic development of communication and language. In the second section, the neural and cognitive mechanisms underlying rhythm processing in humans will be illustrated according to the nowadays models of Cognitive Neuroscience. Lastly, the ontogenetic roots of rhythm perception in the human womb will be described in the third section, exploring the link between prenatal experiences and sensory processing of rhythm throughout life.

2.1 Defining rhythm in music and language

One of the first definitions of rhythm comes from Plato, who defined rhythm as *order in movement*. This definition well describes one of the core components of rhythm: *periodicity*. Periodicity indicates some regular pattern of events varying in time, and it leads to *cycles*, repeating with a particular rate or *frequency*. For instance, seasons alternate periodically, creating a *cycle* that regularly repeats every year. The beginning of every cycle is called *phase*. In physics, an entity producing periodic signals is called an *oscillator*. For instance, a pendulum (*System*, Figure 2.1) is an oscillator generating a periodic signal that can be represented in the time dimension (*Time Series*, Figure 2.1) as a function of amplitude (*Magnitude*, Figure 2.1) over time, or as a Phase Portrait (Figure 2.1) highlighting the cycle's onset and structure. Another important feature of oscillators is that two (or more) entities can synchronize their oscillation patterns, generating a physical phenomenon called *entrainment*.

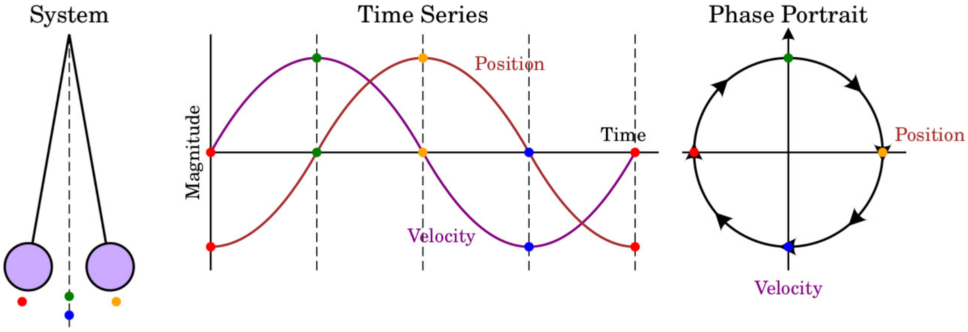


FIGURE 2.1: *System* indicates the pendulum or oscillator; *Time series* indicates the representation of the motion (intensity or Magnitude) over Time, characterized by *Position* and *Velocity*. *Phase Portrait* represents the motion's cycle. From Wikimedia Commons, the free media repository.

Therefore, in its most general definition, rhythm represents a pattern of events repeating regularly over time, and can be shared among entities (McAuley, 2010). Specifically, if all intervals between successive events are equal, then rhythm is defined as *isochronous*. Isochronous rhythms are produced by oscillators that are strictly periodic, as in the case of the beat generated by a metronome. Isochronous rhythms are widespread in biology, including the human body, since heartbeat, respiration, and locomotion all display isochronous patterns (Larsson, 2015; Ravignani and Madison, 2017). Isochrony is also found in the activity pattern of the human nervous system, when synchronous groups of neurons oscillate in the brain (Buzsaki, 2006; Ravignani and Madison, 2017). Even if music and language both feature rich rhythmic organization (Jackendoff and Lerdahl, 2006; Patel and Daniele, 2003), an opening distinction between rhythm in music and language can be outlined exactly about isochrony. That is, isochronous patterns are common in musical rhythms, but this does not occur in the case for human speech. Even if different hypotheses are present in the literature, one classifying languages based on isochrony at the sentence, mora, or syllable levels (Grabe and Low, 2002; Fabb and Halle, 2012) and another stating that speech is an inherently not-isochronous signal (Tuller and Fowler, 1980; Jadoul et al., 2016; Ravignani and Madison, 2017), it is clear that speech rhythm is far away from the strict periodicity of a metronome.

However, isochrony is not a fundamental requirement neither for perceiving nor entraining to a rhythm. Specifically, it is possible to refer both to musical and speech as *quasi-periodic* signals, eliciting rhythm perception and entrainment. In the case of music, when rhythmic patterns are not specifically isochronous, they can be based on a psychologically induced sense of beat (Merker et al., 2009; Ravignani and Madison, 2017). Similarly, the speech signal can vary from a very free rate of conversational interactions to a more periodic structure, as in the case of rhyme and poetry, infant-directed speech, or songs (Obermeier et al., 2016; Kotz et al., 2018). Thus, both music and speech rhythms can display various ranges among a continuum from strictly isochrony to variable, quasi-periodicity (Kotz et al., 2018).

As stated, rhythm perception and entrainment can be elicited by signals at different ranges in the continuum from isochrony to quasi-periodic signals. Moreover, humans display particularly fine abilities in perceiving and producing both quasi-periodic and isochronous patterns across signals and modalities (Fitch, 2015; Motz et al., 2013; Ravignani and Madison, 2017; Ravignani and Norton, 2017). Accordingly, humans tend to regularize isochronous patterns with subtle variations into whole isochronous sequences (Madison and Merker, 2002; Merker et al., 2009). Moreover, speech is perceived as more intelligible when experimentally manipulated by aligning the syllable rate to a heterogeneous rhythmic pattern (Aubanel et al., 2016). Lastly, even infants have been found to i) discriminate between isochronous and non-isochronous patterns at birth (Winkler et al., 2009), ii) synchronize motor patterns to clear periodic signals regardless of subtle variations at 2-4 years of age (Eerola et al., 2006), and iii) display isochronous motor patterns in handwriting during school ages (Pagliarini et al., 2017).

In conclusion, human perception and behavior seem to be attuned to rhythm at various levels. Therefore, one could question the evolutionary purpose of such a peculiar skill. For instance, rhythmic abilities have been hypothesized to have shaped the first proto-interactions between group signals in humans, promoting a turn-taking line that branches in two: one leading to isochrony, as found in music and dance, and the other leading to not-isochronous turn-taking, as found in language and speech (Ravignani and Madison, 2017). This represents just one of the various theories suggesting a deep connection between rhythmic abilities and language development in the human cognitive system. Since what we are is shaped by our history, it is of main interest to explore the evolutionary roots of rhythm and language in order to fully understand the link between these two core abilities. For this reason, the main theories linking rhythm and language abilities in evolution will be briefly illustrated and discussed in the next section .

2.1.1 Evolutionary roots

Throughout history, the ability to process and enjoy musical rhythms has been often considered "a mere incidental peculiarity of the nervous system, with no teleological significance" (cited in Langer, 1942, p. 210; Patel (2006)). However, during the past few decades, new advances on the evolutionary study of musical rhythm processing have been proposed, leading to interesting perspectives on why humans and other animals show rhythm capacities and beat synchronization.

For instance, Merchant and Honing (2014) focused on the phylogenetic continuity of audio-motor abilities among human and non-human primates. As stated, humans display the ability to synchronize, in both period and phase, to complex musical rhythms at variable tempi (Fitch, 2015; Motz et al., 2013; Ravignani and Madison, 2017; Ravignani et al., 2017). Starting from these findings, the authors highlighted how the entrainment to rhythm in humans is a complex cognitive ability, based on the interaction between the auditory and motor systems in the brain (Merchant and Honing, 2014; Grahn and Brett, 2007). Reviewing neuroimaging studies, the authors stated that the motor cortico-basal ganglia-thalamo-cortical circuit (mCBGT) could be the underpinning of such rhythmic behaviors in humans (Harrington et al., 2010; Merchant and Honing, 2014). However, recent findings showed that motor entrainment might not be unique to humans, being rather shared with - at least - some of the other vocal-learning species (Patel et al., 2009; Schachner et al., 2009; Hasegawa et al., 2011). Therefore, the *gradual audio-motor evolution hypothesis* has suggested that rhythmic entrainment might have developed through a gradient of anatomo-functional changes to generate a beat-based mechanism (Merchant and Honing, 2014). Furthermore, since non-human primates do not show this ability, this could represent a first turning point in the evolution of rhythm abilities, occurring roughly 10 million years ago, and going in the direction of continuity between rhythmic and vocal learning abilities across species.

Later, other hypotheses highlighted the ecological aspects of rhythmic signals and processing mechanisms. For instance, Wilson and Cook (2016) remarked how the entrainment of two or more oscillators is a general feature of biological and non-biological physical entities (Figure 2.1). Therefore - they argued - all species could be originally capable of motor entrainment whereas voluntary control or learned coupling of sensory and motor systems might have specifically evolved in some species, according to their evolutionary needs (Wilson and Cook, 2016). Furthermore, the authors listed a few examples of such evolutionary needs, including synchronizing, group locomotion, and breathing to not mask salient environmental sounds (Larsson, 2012) or aquatic animals synchronizing their movements to the frequency of the waves (Wilson and Cook, 2016).

Another reason why, advanced by Ravignani et al. (2014), for the expression of motor entrainment in species is its interactive function. In particular, the authors stated that group behavior is key to understanding the adaptive functions of rhythm, as shown by several studies on chorusing dynamics across species (Ravignani et al., 2014). Despite chorusing behaviors are effectively widespread among taxa (es., *Indomalayan firefly*, Buck (1938); *Neotropical katydid*, Greenfield and Roizen (1993); *Afrotropical frog*, Kaminsky et al. (1999); all reviewed in Ravignani et al. (2014)), humans are the only species showing a spontaneous tendency to engage in group synchrony in the form of a cross-culturally universal propensity to enjoy rhythmic singing and dancing (Merker et al., 2009). Accordingly, possible roots of such complex behaviors have been traced back to our last common ancestor, the African great apes (chimpanzees and gorillas; Fitch (2015), Kotz et al. (2018)). Therefore, complex rhythmic abilities (e.g., sensory-motor entrainment at variable tempi) have accordingly been proposed as an indicative trait of our species, along with language (Merker et al., 2009).

In conclusion, the evolutionary path leading to the complex rhythmic abilities of human species develops hand in hand with sensory, motor, and communicative phylogenetic pressures. Starting from these findings, Patel (2006, 2021) advanced the *vocal learning and rhythm synchronization hypothesis*, trying to identify the last step of the evolutionary path into the neural circuitry for vocal learning. This hypothesis has become the most discussed and tested to date. For instance, a recently published paper showed evidence of beat synchronization in rats, a non vocal learning species (Ito et al., 2022).

Therefore, although further research is needed to clarify the extent to which high rhythmic abilities might be considered as specie-specific, discussing the *vocal learning and rhythm synchronization hypothesis* is crucial for our understanding of the link between rhythm and language development. Therefore, it will be illustrated in the next and last paragraph of this section.

2.1.2 The revised vocal learning and rhythm synchronization hypothesis

As mentioned in the previous paragraph, the *vocal learning and rhythm synchronization hypothesis* (VLH, Patel (2006)) originally posits that the entrainment of motor patterns to an external rhythm, defined as "beat perception and synchronization" (BPS), relies on fine audio-motor forebrain connections originally evolved to sustain complex vocal learning. Complex vocal learning is defined as the animal's need for auditory input to develop its typical species-specific vocalizations, since the auditory input creates a template that shapes the acquisition of the animal's own vocal productions (Patel, 2006). Only a narrowed set of mammals show complex vocal learning (i.e., cetaceans, pinnipeds, and humans) and three groups of birds (i.e, songbirds, parrots, and hummingbird; Tyack (2020); Patel (2006)). When A. Patel (2006) first outlined the hypothesis, BPS was not found in any other animals except humans (Bispham, 2006). However, in the following 15 years, evidence showing signs of BPS in parrots (Patel et al. (2009); Schachner et al. (2009); Hasegawa et al. (2011), Figure 2.2) partially supported the hypothesis. Therefore, Patel (2021) revised his hypothesis by addressing evidence on parrots. In the revised VLH (rVLH), vocal learning is better classified as a continuum, encompassing various abilities that may be evolved for different purposes across species (Patel, 2021; Petkov and Jarvis, 2012; Arriaga et al., 2012). Specifically, vocal plasticity (i.e., the ability to modify vocal repertoire based on social auditory experiences) is core for the rVLH, since it requires fine audio-motor integration in the forebrain, a foundational feature for BPS. Crucially, vocal plasticity is higher in humans and parrots, who are classified together as high vocal learners in the vocal learning continuum (Figure 2.2).

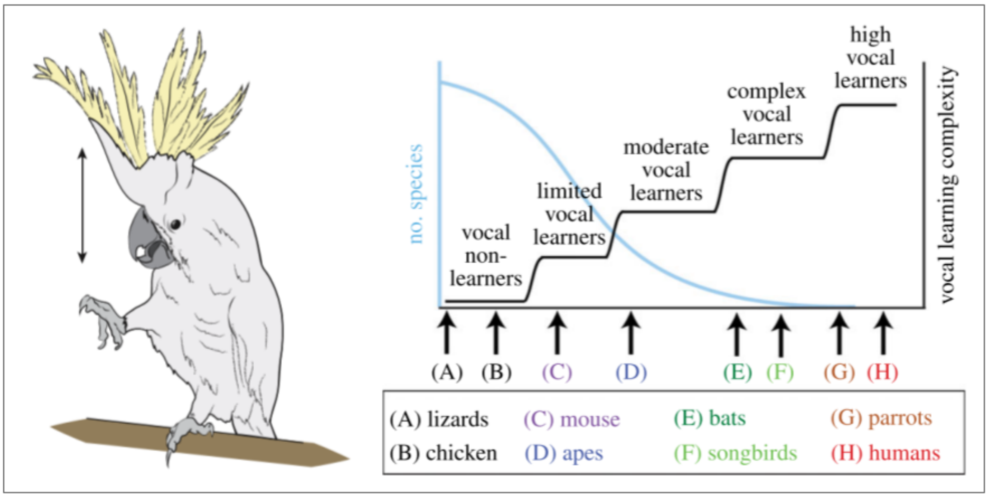


FIGURE 2.2: Snowball the cockatoo (left) is capable of entraining to a musical stimulus and varying the head movements as the tempo of the music is experimentally changed (Kotz et al., 2018). The vocal learning continuum hypothesis (right), with step-wise continuous ability of vocal learning among vertebrates (right y-axis), from simple to more complex forms (x-axis). As vocal learning complexity increases, there are a decreasing number of species with the ability (left y-axis). (A–H) Proposed example species at each step on the continuum (Patel, 2021).

Therefore, the rVLH proposes that high vocal learning is foundational to spontaneous, sporadic BPS to rhythm. Since parrots and humans are the only species (to date, e.g. Ito et al. (2022)) known to spontaneously engage in BPS, rVLH states that high vocal learning is a necessary but not sufficient prerequisite for spontaneous BPS. Moreover, precursors of both BPS and high vocal learning can be found in human ancestors that already showed rhythmic social behaviors. The evolution of such complex traits might then be the by-product of gene-culture coevolution. Accordingly, parrots and humans both share the ability to imitate body movements and live in social environments. Therefore, the author argues that early humans might have initially sporadically used BPS for social behaviors, for instance to share cooperation and social connections, and since these actions were advantageous for sexual and survival purposes, BPS might then have specialized via gene-culture evolution. This means that culture (e.g., imitation of successful behaviors) and genetic dynamics (e.g., the transmission of genetic variants enhancing the ability of sustained BPS; Patel (2021); Gordon et al. (2021)) might have driven the evolutionary line to where it is today.

To summarize, the evolutionary roots of rhythmic abilities can be traced back from the capacity to produce periodic motions (nearly universal among animals) and beat perception (shared with non-human primates), to the ability to motor entrain with conspecifics (shared with animals in chorusing behaviors) and to external rhythms (BPS; which seems to be shared only within high vocal learner species). This large corpus of evidence therefore suggests that complex rhythmic and linguistic abilities might evolved hand in hand with vocal learning, resulting in specialized cognitive processes and underlying neural networks for music and language processing in humans. However, further research is needed to include the growing evidence of rhythmic abilities in other species to a coherent framework (e.g., Ito et al. (2022)). This might significantly contribute to our understanding of the links between rhythm in music and language since it allows us to formulate testable predictions not only within the evolutionary and comparative fields but also in neuroscience in general and language and music development in particular. Accordingly, the phylogenetic history of complex abilities might be informative about the functional adaptations that led to the neural and cognitive specialization through continue interactions with the external environment. Accordingly, this line of literature brought significant knowledge about which cognitive and neural mechanisms are essential to and are shared between language and musical rhythm processing, as illustrated in the next section.

2.2 Neural and cognitive rhythm processing

2.2.1 The neuroscience of rhythm

Nowadays research has well described the sophisticated cortico-subcortical network serving both temporal and rhythm processing in the human brain. Specifically, the ability to produce and perceive rhythms seems to rely on a set of cortical neural circuits, including the prefrontal cortex (PFC), supplementary motor area (SMA), thalamus, basal ganglia, dorsal premotor cortex (dPMC), posterior superior temporal gyrus/middle temporal gyrus (pSTG/MTG) and parietal regions near the angular gyrus (AG), illustrated in Figure 2.3 and described below.

Connections in orange and light blue in Figure 2.3 indicate those regions sustaining BPS, of particular interest for language in humans (Patel, 2006, 2021; Gierhan, 2013). Specifically, orange connections link secondary auditory regions in the posterior superior temporal gyrus/middle temporal gyrus (pSTG/MTG) to parietal regions near the angular gyrus (AG), while light blue connections link regions near the angular gyrus to the dorsal premotor cortex (dPMC). These connections correspond to two branches of the superior longitudinal fasciculus (SLF): the temporo-parietal branch (SLF-tp) and the second branch (SLF II). Further functional imaging studies in humans have revealed that the motor cortico-basal ganglia-thalamo-cortical circuit (mCBGT; dark blue in Figure 2.3) is also involved in sequential (Grafton et al., 1995) and temporal (Harrington et al., 2010) processing, as well as on rhythmic behaviors (Grahn and Brett, 2007; Merchant and Honing, 2014). Lastly, the prefrontal cortex (PFC, red line in Figure 2.3) is also involved in rhythm processing, playing its role in attention-dependent temporal processing (Grahn, 2012; Schwartze and Kotz, 2013).

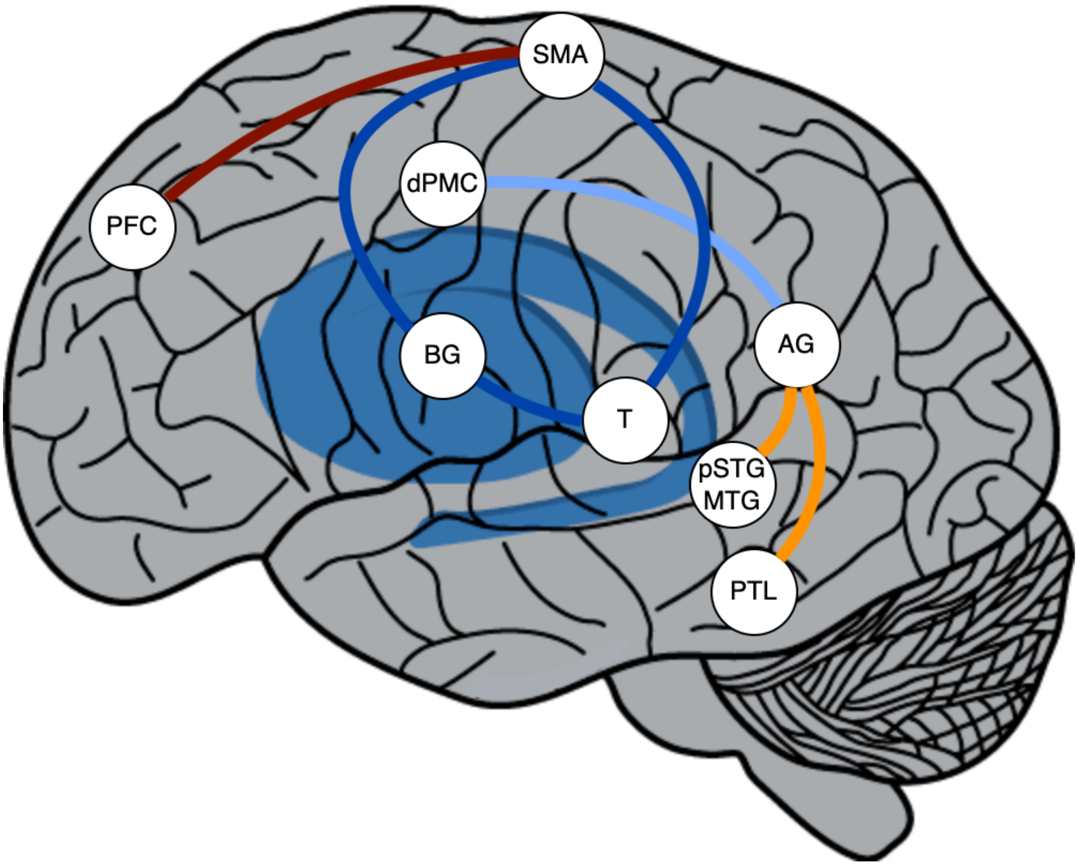


FIGURE 2.3: Prefrontal cortex (PFC), supplementary motor area (SMA), thalamus (T), basal ganglia (BG), dorsal premotor cortex (dPMC), posterior superior temporal gyrus/middle temporal gyrus (pSTG/MTG), angular gyrus (AG), and posterior temporal lobe (PTL). In orange and light blue, networks of particular interest for beat perception and synchronization (Patel, 2021). In dark blue, the motor cortico-basal ganglia-thalamo-cortical circuit (Merchant and Honing, 2014). In red, attentional circuits involving frontal regions (Kotz et al., 2018). Original brain image (gray) by Casey L. Henley.

.

Therefore, the neural circuits underlying rhythm perception and production in the human brain involve cortical and subcortical areas traditionally associated with motor (i.e., basal ganglia, SMA, dPMC) and linguistic (i.e., pSTG/MTG, AG, PTL, dPMC) functions. These results further support the existence of a deep link between rhythm and language, and the important role of motor processes in such relationship at the brain level.

Moving to the cognitive aspects, a large corpus of evidence shows a substantial overlap between the rhythm processing in speech and music. These results led to the formulation of several theories aimed at characterizing the shared processing of music and speech rhythms. However, every theory mainly focused only on a specific mechanism supporting rhythm in one domain. To unify the evidence into a coherent model, a parsimonious framework has been recently proposed: the *processing rhythm in speech and music* (i.e., PRISM) framework (Fiveash et al., 2021), illustrated in the next paragraph.

2.2.2 The *processing rhythm in speech and music* framework

The recently formulated PRISM framework (Fiveash et al., 2021) proposes that: i) precise, fine-grained auditory processing, ii) synchronization / entrainment of neural oscillations, and iii) sensorimotor coupling are the three critical elements underlying rhythm processing in music and speech (see Figure 2.4). As stated, this framework was meant to unify the evidence that emerges separately across theories from different research fields. Previously proposed approaches included theories on shared elements of music and speech rhythms (Fujii and Wan, 2014; Tierney and Kraus, 2014; Goswami, 2011, i.e., the sound envelope processing and synchronization and entrainment to pulse hypothesis, the precise auditory timing hypothesis, and the temporal sampling framework for developmental dyslexia); more broad theories on music processing, such as the OPERA hypothesis, suggesting that Overlap, Precision, Emotion, Repetition, and Attention drive the influence of music training on speech processing (Patel, 2011); and lastly, theories involving the sensorimotor system (Patel and Iversen, 2014; Jones, 2018; Friston, 2010, e.g., action simulation for auditory prediction, active sensing, the dynamic attending theory, and predictive coding) are also part of the formulation of the current framework.

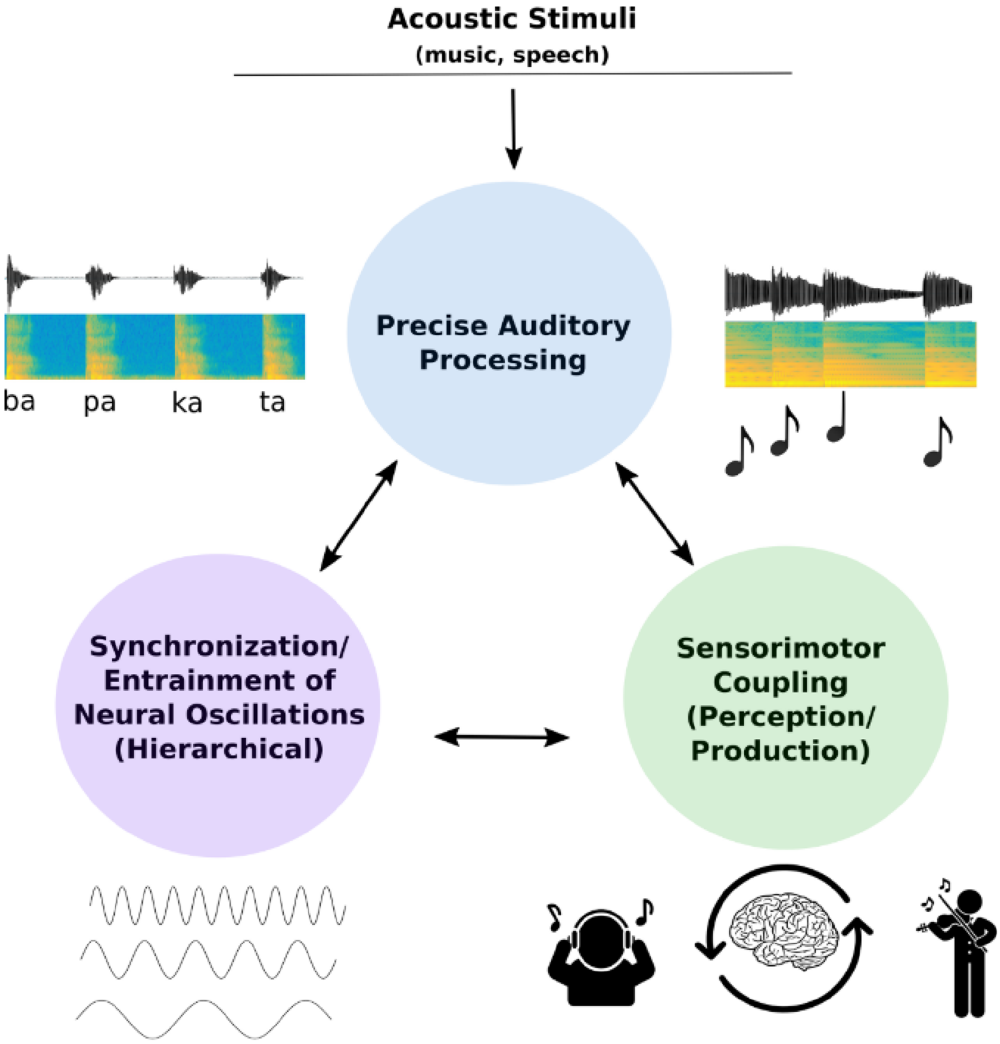


FIGURE 2.4: The three mechanisms underlying rhythm processing in music and speech observed across different theories: precise auditory processing; synchronization/entrainment of neural oscillations to external rhythm; and sensorimotor coupling. From Fiveash et al. (2021).

Precise auditory processing

Precise auditory processing refers to the ability to discriminate subtle variations in timing, pitch, and timbre (Kraus and Chandrasekaran, 2010). This ability is crucial to perceive subtle changes in speech (e.g., /ba/ from /pa/) and music on the millisecond level (Patel, 2011). Sensorimotor coupling and entrainment of neural oscillations also interact with auditory processing (Repp and Penel, 2004; Giraud and Poeppel, 2012). The result is a set of complex auditory processing abilities that allows it to process both music and speech signals, with the possibility to transfer improvements from one domain to the other. Specifically, it has been proposed that music training might enhance speech processing thanks to the overlapping brain networks supporting general timing abilities, necessary for processing music and speech rhythms (Patel, 2011). Consistently, studies report that music training enhances the neural response (the mismatch negativity, MMN) to violations of temporal structure in both music and speech (Zhao and Kuhl, 2016) as well as to syllable duration and vowel onset time deviants (Chobert et al., 2014), resulting in increased speech segmentation skills (François et al., 2013). This evidence therefore suggests that music rhythm training can improve speech rhythm processing. Furthermore, the regularity of music rhythms has also been found to improve the precision of the auditory processing and neural entrainment abilities, positively impacting phonological skills and the neural tracking of the speech envelope (Flaugnacco et al., 2015; Goswami, 2011). Therefore, both music and speech processing rely on precise auditory timing which is also linked to neural entrainment and sensorimotor skills.

Synchronization and Entrainment of Neural Oscillation

Neural oscillations are inhibitory and excitatory electrical patterns recurring regularly and ubiquitously in the brain (György Buzsáki, 2019). Neural oscillations entrain to auditory rhythms underlying both music and speech perception (Fujioka et al., 2012; Kösem and Van Wassenhove, 2017). The entrainment of neural oscillations to rhythms has been linked to temporal attention (Jones et al., 2002), prediction (Arnal and Giraud, 2012), hierarchical processing (Poeppel and Assaneo, 2020), and communication between brain regions (Assaneo and Poeppel, 2018; Fiveash et al., 2021).

Specifically, endogenous oscillatory activity in the brain entrains in phase to rhythm and this sustains the temporal allocation of attention toward recurrent (*predictable*) portions of the signal (Jones, 2018). Accordingly, several behavioral and electrophysiological studies report the beneficial effects of neural entrainment on perception, learning, and memory (Hickey and Race, 2021; Henry et al., 2014; Barnes and Jones, 2000). Furthermore, neural entrainment to music rhythm was found to persist and extend to subsequent speech, with effects at the behavioral (Cason et al., 2015a) and neural level (Cason and Schön, 2012). Neural entrainment was also found to occur at multiple hierarchical levels in music and speech. Accordingly, when listening to music, different beat- and meter-based frequencies of neural activity are registered (Fiveash et al., 2020; Nozaradan et al., 2012). Similarly, delta (1-3 Hz), theta (4-8 Hz), and low gamma (25-35 Hz) frequency bands respectively characterize the neural response to the speech at the phrasal, syllabic, and phonemic level (Giraud and Poeppel, 2012; Ghitza, 2011). Lastly, neural oscillations have been found to be associated also with precise auditory processing (Goswami, 2011; Poeppel, 2003) and sensorimotor coupling (Morillon and Baillet, 2017; Fiveash et al., 2021). Therefore, the entrainment of endogenous oscillations to external rhythms has been shown to play a central role in music and speech processing. Furthermore, the regular rhythmic structure (at different hierarchical levels) and temporal precision of music represents an ideal stimulus to elicit neural entrainment, with potential benefits for speech processing.

Sensorimotor coupling

Sensorimotor coupling refers to the connections between sensory and motor cortices, and plays a central role in the perception and production of rhythm in music and speech. Specifically, studies report neural activity in the motor cortex (supplementary motor area or SMA, pre-SMA, and premotor cortex) when listening to rhythmic music or speech patterns (Chen et al., 2008; Gordon et al., 2018; Möttönen et al., 2013). Indeed, the motor system is involved in rhythm perception and the auditory system is involved in rhythm production, suggesting a link between perception and production in the two domains (Guenther and Hickok, 2015; Hickok et al., 2011). Moreover, sensorimotor coupling seems to sustain language acquisition with higher grades of white matter pathways between frontal and auditory regions enhancing word learning (Assaneo et al., 2019; Bruderer et al., 2015).

Interestingly, motor regions (premotor cortex, the SMA, pre-SMA and the putamen) have been hypothesized to i) receive sensory input from the auditory cortex, ii) create motor planning based on this input, and iii) send timing prediction back to the auditory cortex (Patel and Iversen, 2014; Cannon and Patel, 2021). Neural oscillations are also involved in this process and specifically seem to play a role in the connection between regions, amplifying the sensory input at predicted times (Morillon et al., 2015; Schroeder et al., 2008). Therefore, both auditory and motor regions participate in the rhyhtmic speech and music processing, with sensorimotor coupling specifically sustaining temporal regularities across domains.

Clearly, auditory input in rhythm perception has been extensively explored across studies; however, hearing is not the unique sensory modality involved in this process (Figure 2.5). Specifically, first experiences with rhythm, occurring already in the womb, encompass multiple sources of rhythmic stimulation, including vestibular, tactile, and somatosensory inputs produced by maternal body's movements (Lecanuet and Schaal, 2002; Provasi et al., 2014). Consequently, vestibular and somatosensory input emerged as core aspects of rhythm perception after birth and during development (Trainor et al., 2009; Phillips-Silver and Trainor, 2005; Tichko et al., 2021). Therefore, although most studies on rhythm perception are biased toward auditory versus vestibular, tactile, or somatosensory stimulation (Provasi et al., 2014; Bremner and Spence, 2017), for their centrality in rhythm and language development, vestibular and somatosensory aspects of rhythm perception will be discussed in the next and last section of this Chapter.

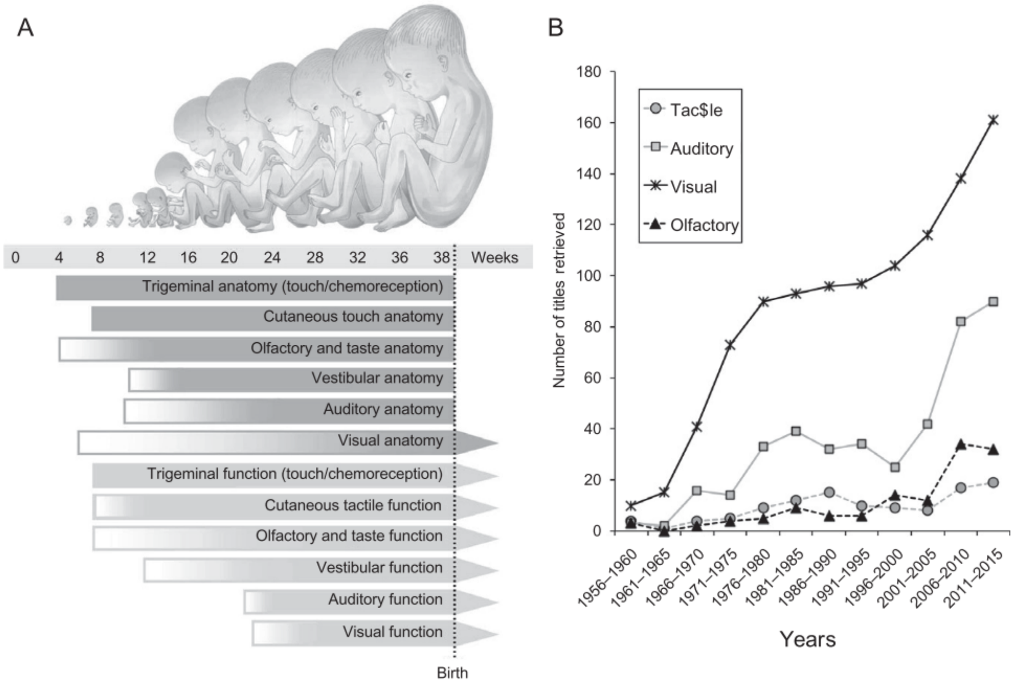


FIGURE 2.5: (A) The emergence of the anatomy and function of multiple sensory systems during human gestation. Dark gray bars indicate the emergence and maturation of the senses (usually provided by histological evidence). Shading of the dark gray bars indicates the time between the first observation of the sensory apparatus and its full anatomical maturation. Light gray bars indicate the onset of function of the senses. Shading of the light gray bars indicates the uncertainty in the literature concerning the first age at which function is observable. Arrows pointing beyond birth indicate continued postnatal development. (B) The number of articles retrieved by PsycINFO for searches on sensory development (y-axis) was restricted to 5-year time windows from 1956 to 2015 (x-axis). From Bremner and Spence (2017).

2.3 Feeling rhythm: VTS rhythm perception

The human body is capable of perceiving tactile input through cutaneous sensation, interoception, and proprioception (Gallace and Spence, 2014). Interoception provides sensations (typically unconscious) about organ functioning, homeostasis, digestion, and respiration (Craig, 2009); proprioception informs about how the body and limbs are placed or moving in space, shaping the body and limbs' representation; cutaneous stimulation is integrated with interoception and proprioception to interpret our tactile environment through haptics (Bremner and Spence, 2017). All these signals are processed in the primary somatosensory cortex, giving rise to somatosensory experience (Penfield and Rasmussen, 1950). A further sense of body position, velocity, and duration of motion come from vestibular input, which originates in the labyrinth, is refined in the brainstem and cerebellar circuits, and then is relayed to the cortex (Choi et al., 2021). All these components together play a role in rhythm perception and will be defined from now on as vestibular-tactile-somatosensory (VTS) experience (Provasi et al., 2014).

2.3.1 Prenatal rhythm perception

As stated in the previous section, early experiences with rhythm begin in the womb, where fetuses are exposed to different sources of stimulation, including VTS input produced by the maternal body (Lecanuet and Schaal, 2002). As shown in Figure 2.5, the VTS system is the first to develop in the human fetus: cutaneous and trigeminal somatosensory receptors mature at around 4–7 weeks (Humphrey, 1964). Consistently, while the early emergence of somatosensory anatomy and function is shortly followed by vestibular and chemosensory functioning, auditory and visual functioning are first observed much later in prenatal development (Bremner and Spence, 2017). Accordingly, from the third trimester of gestation fetuses can perceive rhythmic signals from the intrauterine and extrauterine environment through bone conduction (Sohmer et al., 2001). Specifically, the maternal voice is transmitted directly to the amniotic fluid via body tissues and bones, with F0 and the first overtones being fully conducted through the spine and the pelvic arch (Petitjean, 1989).

External low frequencies (corresponding to the first four octaves of the piano) are well perceived *in utero*, with bone conduction allowing for the intelligibility of both speech and music rhythmic patterns (Granier-Deferre et al., 2011). Therefore, 'hearing' occurs through VTS mechanisms, and early rhythmic experiences are mainly provided through this sensory modality.

Interestingly, near-term fetuses can differentiate between VTS rhythms, for instance displaying changes in heart rate in response to different mother's rhythmic movements (Cito et al., 2005). Crucially, several studies proposed that the early VTS/auditory experiences with rhythms, resulting from the intra/extrauterine environment (Ullal-Gupta et al., 2013; Teie, 2016) and maternal movements (Provasi et al., 2014; Rocha et al., 2021), might shape the developmental trajectories of rhythmic abilities after birth. Consistently, premature infants who receive VTS stimulation increase their respiratory rate compared to infants receiving no stimulation, adapting their breathing to the rate and acceleration of VTS stimuli (Zimmerman and Barlow, 2012). Moreover, caregivers from all cultures use VTS rhythms to calm their newborns, for instance by moving them back and forth or by rocking them while singing or walking (Provasi et al., 2014). Accordingly, a recent study demonstrated that the cadence of parent movements and walking is related to the spontaneous motor tempo of young infants, demonstrating that exposure to passive rhythmic actions (experienced both prior and shortly after birth) shapes first rhythmic behaviors in infancy (Rocha et al., 2021). Prior to this, the seminal work from Phillips-Silver and Trainor (2005) showed that 7-months-old infants who were bounced in synchrony to a given rhythm preferred to listen to a matching auditory stimulus compared to a mismatching one. Therefore, the authors concluded that passive movements generating VTS stimulation influence the perception of auditory rhythm in infants (Phillips-Silver and Trainor, 2005). After this work, the research investigating VTS rhythmic stimulation kept growing with further evidence coming from adult participants leading to the nowadays dynamical, embodied, and ecological approaches to rhythm processing, as described in the next section.

2.3.2 Rhythm in the body

Evidence on the deep connections between rhythm and movement come from studies investigating the similarity between locomotion and preferred tempo (Todd and Lee, 2015), the human ability to synchronize motion to rhythms (Repp and Doggett, 2007), and from neuroimaging studies showing brain regions responsible for synchronized movement being modulated by auditory metrical structure (Haueisen and Knösche, 2001; Lahav et al., 2007). Moreover, rocking infants to a rhythm has been found to shape their listening preference for auditory stimuli (Phillips-Silver and Trainor, 2005) with VTS/auditory rhythmic abilities pointed as a potential scaffold for early interpersonal synchrony between mothers and infants (Trehub, 2003). Similar findings come from studies with adult participants, showing that body movement can bias the auditory encoding of ambiguous rhythms (Phillips-Silver and Trainor, 2007). Interestingly, watching someone else moving in time with a meter is not sufficient to elicit similar effects, demonstrating that the active movement of the body is crucial in rhythm encoding and processing (Phillips-Silver and Trainor, 2007). The authors then tested whether the passive movement was associated with VTS-to-auditory effects on rhythm perception in adults as in infants (Phillips-Silver and Trainor, 2008). Surprisingly, it was found not only that passive VTS motion is sufficient to bias auditory rhythm encoding but also that vestibular stimulation alone is sufficient to elicit such effects (Phillips-Silver and Trainor, 2008; Trainor et al., 2009). Specifically, the passive motion of the head (but not of the legs) as well as a galvanic stimulation delivered to the vestibular nerve both were found to drive rhythm perception in adult participants (Phillips-Silver and Trainor, 2008; Trainor et al., 2009). Therefore, this evidence further confirms the crucial role of VTS input in perceiving and processing rhythmic patterns.

Findings on VTS stimulation affecting rhythm processing in adults are of particular interest since, at this stage of development, individuals are quite distant from their prenatal experience with VTS rhythms as perceived in the womb. However, evidence suggests that active and passive motion generating vestibular, tactile, and somatosensory experience not only shapes the early development of structural and functional mechanisms underlying rhythm processing, but also influences rhythm processing during the lifespan. This can be explained by considering that, in the adult's as in the infant's brain, the encoding of VTS input conveyed by physical proximity to a sound source consists of the same energy (vibratory) used to encode auditory stimuli (Ammirante et al., 2016); moreover, mechanoreceptors and ear cells are similarly structured and comparable in response characteristics such as, for instance, the loudness summation of tones closely spaced in frequency (Marks, 1979; Ammirante et al., 2016). Consistently, VTS inputs have been found to activate the auditory cortex (Caetano and Jousmäki, 2006; Ammirante et al., 2016), with VTS and auditory inputs being confused when presented simultaneously or in alternation (Von Bekesy, 1959; Gescheider and Niblette, 1967)). Together, these findings suggest that vibrations might evoke comparable low-level responses between VTS and auditory modalities (Ammirante et al., 2016).

Accordingly, not only psychophysical experiments demonstrate high-level interactions between audio and tactile sensory systems, but neuroimaging studies further suggest that these interactions also occur in early sensory areas (Crommett et al., 2017, 2019; Fery et al., 2021; Caetano and Jousmäki, 2006). Therefore, this might contribute to explaining the recent findings showing that i) the perception threshold curves of audio and haptic rhythmic gradients are the same (Bernard et al., 2022), ii) that the accuracy in synchronizing to a tactile metronome can equal that of an auditory metronome (Ammirante et al., 2016), and iii) that the discrimination of complex vibrotactile stimuli involves the cortical integration of spectral information filtered through frequency-tuned skin receptors (Ammirante et al., 2013).

The *dynamical, radically embodied, and ecological theory* of rhythm

Based on previous findings on VTS-auditory coupling, Tichko et al. (2022) specified a nonlinear, dynamical system in which two oscillatory neural networks, representing the auditory and motor systems, interact through weak, non-specific coupling. By doing so, the authors demonstrated that this auditory-motor coupling is responsible for bootstrapping short-term Hebbian plasticity in the auditory system, explaining the VTS effects on infant and adult rhythm encoding in terms of enhanced vestibular-related frequencies in the auditory network activity (Tichko et al., 2022). Tichko et al. (2022) therefore conceptualized a dynamical, radically embodied, and ecological theory of rhythm by grounding on the Neural Resonance Theory (Large and Snyder, 2009; Large et al., 2015). Specifically, this theoretical framework proposes that ontogenetic changes in rhythm perception and action occur via the resonance and the attunement of coupled auditory-motor systems and rhythmic inputs across development (Tichko et al., 2022). That is, the coupling of auditory, motor, and vestibular systems during development is explained by means of two well-established properties of neurobiological systems: neural oscillation (i.e., cyclic patterns of neural activity thought to support rhythm perception-action during infancy and adulthood) and Hebbian plasticity (i.e., changes in synaptic strength based on the co-activation of neural units) (Large and Snyder, 2009; Nozaradan et al., 2011; Munakata and Pfaffly, 2004). Therefore, the coupling between auditory and motor systems, which is hypothesized to reflect diffuse connectivity across sensorimotor systems during early development, can account for VTS influences on rhythm perception.

Interestingly, the VTS-auditory coupling has been also found to sustain language acquisition across development. For instance, synchronous tactile cues were shown to help 4- and 5-month-olds to find words in continuous speech, promoting word learning (Seidl et al., 2015; Abu-Zhaya et al., 2017). Moreover, audio-tactile stimulation resulted in enhanced ERPs and higher beta-band activity, compared to auditory-only stimulation, in 8-months-old infants (Tanaka et al., 2018).

Consistently, parents tend to spontaneously synchronize the location and timing of their touches on the infant's body with word rhythm in infant-directed speech (Tincoff et al., 2019; Custode and Tamis-LeMonda, 2020; Lew-Williams et al., 2019). Later in development, preschoolers who can entrain their motion to an external beat have been found to display more faithful neural encoding of temporal modulations in speech and higher scores on tests of early language skills (Woodruff Carr et al., 2014). By contrast, rhythmic deficits are shared by different atypically developing populations, including developmental dyslexia (Beker et al., 2021), attention deficit disorder (Carrer, 2015), autism (Franich et al., 2021), and developmental coordination disorder (Chang et al., 2021).

In summary, rhythmic abilities are deeply rooted in the human brain. They originated from our need for social interaction and bonding and have evolved hand-to-hand with communication and language abilities across ages. In ontogenetic evolution, first experiences with rhythm occur already in the womb, with different sources of rhythmic stimuli being available to the human fetus. Among all the sensory modalities, VTS input seems to play a crucial role in rhythm processing, with neural coupling dynamics allowing for VTS influences on auditory encoding. Therefore, this Chapter demonstrated the link between rhythm and language abilities through phylogenetic evolution (Section 2.1); the neural and cognitive underpinnings evolved in humans for rhythm perception in speech and music (Section 2.2); and the ontogenetic roots for rhythm processing in the womb (Section 2.3). Lastly, evidence suggests that rhythmic abilities might have important impacts on typical and atypical language development. Specifically, this will be the starting point for the next Chapter, investigating the role of rhythm in typical and atypical language development across the lifespan.

Chapter 3

Rhythm in language acquisition

3.1 Linguistic rhythm

3.1.1 Definition

Linguistic rhythm can be defined as the alternation of weak and strong elements at different levels of the prosodic hierarchy (Nespor and Vogel, 2007; Langus et al., 2017). Specifically, rhythm arises from the alternation between consonant and vowel at the segmental level (Ramus et al., 1999), between stressed and unstressed syllables at the metrical foot level (Hayes, 1995), and between stressed and unstressed words at the phonological phrase level (Nespor et al., 2008).

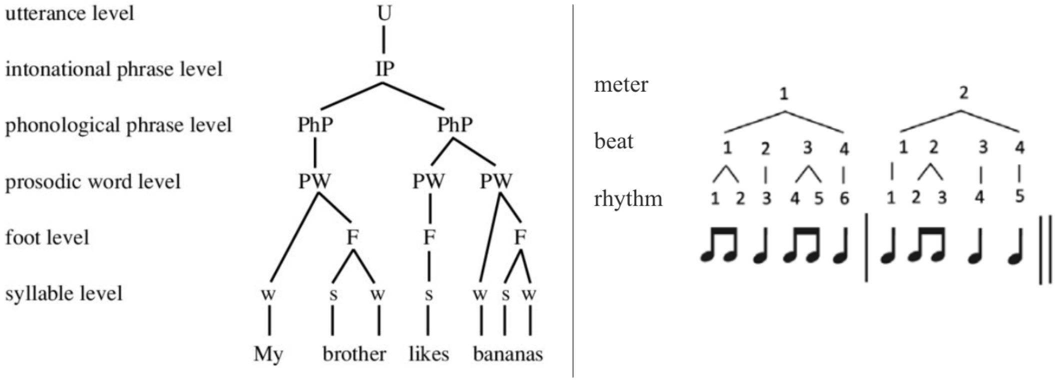


FIGURE 3.1: The speech line (left) "My brother likes bananas" is characterized by the syllable level (w is for weak and s for strong syllables), the foot level (F indicates the feet units), the prosodic word level (PW indicates the prosodic word units), the phonological phrase level (PhP indicates the phonological phrase units), the intonation phrase level (IP indicates the intonational phrase unit), and the utterance level (U indicates the utterance unit). The music line (right) is grouped into rhythms, beats, and meters. Adapted from Ladányi et al. (2020); Gerken and McGregor (1998).

As shown in Figure 3.1, language - as music - is a hierarchical signal. Specifically, its prosodic hierarchy has a nested structure with all constituents included into higher ones (Selkirk, 1984). In this structure, metrical feet never exceed the prosodic word boundaries, and the prosodic words never exceed phonological phrases (Selkirk, 1984; Langus et al., 2017). Similarly, stress at the prosodic word level is instantiated over stressed syllables at the metrical feet level; therefore, stress at higher levels of the prosodic hierarchy always matches with the stress position in one of the lower levels. Importantly, relative stress (or *prominence*) in the prosodic hierarchy is always generated by perceivable changes in pitch, intensity, or duration (Cutler et al., 1997; Lehiste, 1970; Langus et al., 2012). Specifically, humans tend to group elements alternating in prominence - being this generated through pitch/intensity or duration - following the iambic-trochaic law (Bolton, 1894; Nespor et al., 2008). Accordingly, units varying in duration are grouped into iambic patterns (i.e., weak-strong / short-long) whereas units varying in pitch or intensity are grouped as trochees (i.e., strong-weak / high-low).

The conceptualization of speech rhythm in terms of prominence, grouping, and stress rather than the historical categorization into stress-timed and syllable-timed languages, allows for a better understanding of rhythm processing across different languages, and across linguistic as well as music signals (Arvaniti, 2009; Ding et al., 2017). Indeed, the complementary view of rhythm in music and language frames rhythm as a perceptual phenomenon arising from hierarchically structured signals. Specifically, the hierarchical structure of both music and speech is centered around the concept of strong and weak elements, with strong units providing a perceptual framework in which weak units are nested. Therefore, rhythm in language - as in music - is essentially a hierarchical patterning of events. Furthermore, this framework allows researchers to overcome the limitations of considering rhythm as an exclusive feature of isochronous signals (see Section 2.1). Indeed, even though speech rhythm is not isochronous, the alteration of strong and weak elements at multiple levels of the prosodic hierarchy is regular and predictable, leading to expectations on the upcoming signal (Peelle and Davis, 2012; Giraud and Poeppel, 2012; Arvaniti, 2009).

The ability to generate expectations based on patterns of prominence and stress has also parallels with music, where the expressive variations in timing can be anticipated by the structure that arises from beat and metric patterning (Grahn, 2012; Kotz et al., 2018). As discussed in Section 2.1, isochrony has been found to be not a fundamental requirement neither for perceiving or entraining to a rhythm, with music and speech being better defined as quasi-periodic signals capable of generating both rhythm perception and entrainment (Merker et al., 2009; Kotz et al., 2018).

In summary, linguistic rhythm can be defined as 'the product of prominence and patterning' and important parallels can be made between this structure in language and music (Arvaniti, 2009, p.61). Moreover, the human mind spontaneously tends to perceive the emergent structure of rhythm at different hierarchical levels and to formulate expectations on this base. Therefore, researchers hypothesized that the rhythmic structure of speech might serve as a scaffold to language development, bootstrapping the first stages of language acquisition and the later development of complex linguistic abilities, as illustrated in the next paragraph.

3.1.2 Language acquisition

In Chapter 2, the ontogenetic roots of rhythm processing were traced to the rhythmic stimuli perceived by human fetuses in the womb. Accordingly, external and maternal voices are available in the intrauterine environment. In particular, extra-uterine voices are perceived as low pass filtered by maternal tissues and the amniotic fluid (DeCasper and Fifer, 1980; DeCasper and Spence, 1986; DeCasper et al., 1994). This filter action only left unaltered the prosodic component of speech including the rhythmic structure (Cooper and Aslin, 1990). Therefore, the rhythmic structure of language is present since the first experiences with this complex signal. Furthermore, evidence of perceptual learning at this early-stage lead to the conclusion that the human cognitive system is shaped by auditory and vestibular experiences with regular patterns of speech occurring already in the uterine environment (Ullal-Gupta et al., 2013; Granier-Deferre and Busnel, 2011, see Chapter 2).

Accordingly, signs of previous experience with native speech rhythms can be found in the newborn ability to distinguish native from foreign languages (Mehler et al., 1988). Specifically, such ability is mainly based on prosodic features since it encompasses low-pass sentences in which all phonetic content was removed and only rhythm and contour were available. Moreover, newborns can also discriminate between languages based on different rhythmic structures (e.g., English and Dutch) but they fail to tease apart languages sharing similar rhythmic features (Nazzi et al., 1998, e.g., English and Italian). Therefore, rhythm is a crucial component in the earliest representations of language.

Growing into infancy, babies start discovering and segmenting individual linguistic units (i.e., words) from continuous speech. Rhythm is a highly informative cue to word boundaries, since the way in which prominence arises from strong and weak syllables alternating at the metrical feet level is usually uniform across words and peculiar to languages. For instance, prominence in English is conveyed by increasing pitch/intensity (Cutler, 1984) and almost 90% of words in this language receive prominence on the first syllable composing a trochaic stress pattern (Carlson, 1989). Accordingly, English-learning infants recognize and prefer words displaying such rhythmic structure from 9 months (Jusczyk et al., 1993). Moreover, they can extract trochaic words from continuous speech and tend to treat strong syllables as word onset (Jusczyk et al., 1999). By contrast, French-learning infants do not process trochaic rhythms in the same way, failing to segment words with this pattern from continuous speech (Polka et al., 2002). This can be attributed to the rhythmic structure of the French language, in which prominence is realized with longer duration and thus, iambic patterns easily arise (Cutler, 1984). Consistently, French-learning infants rely on iambic rhythmic patterns in identifying words (Hallé and de Boysson-Bardies, 1996). Therefore, rhythm is functional to the emerging segmentation abilities in infancy.

While infants develop the ability to identify linguistic units in the speech stream, they start to access word-level representations. In a seminal study, Morgan and Saffran (1995) demonstrated the central role of rhythm in driving this process. In this study, infants were presented with sequentially recurring syllables with different rhythmic patterns as well as non-sequential stimuli. At 6 months of age, they were found to create units only based on rhythmic regularities, not following the sequential, phonological grouping strategy; at 9 months of age, infants grouped regular units based on both rhythmic and segmental structures, disregarding the sequences displaying a contradictory rhythm. Therefore, the authors argued that young infants strongly rely on rhythmic regularities while older infants combine rhythmic and segmental cues to build word-like representations.

The main gain from word representation is lexical access. Rhythm has been found to sustain infants in this developmental challenge too. Indeed, they rely on the rhythmic structure of their language to predict and allocate attentional resources toward highly informative portions of speech (Giraud and Poeppel, 2012; Arvaniti, 2009). For instance, English- and Dutch-learning infants recognize familiar, disyllabic words even when mispronunciations occur on weak syllables (e.g., *dirty* as *dirny*), whereas mispronunciations that occur on strong syllables (e.g., *dirty* as *nirty*) interfere with the recognition of a familiar word (Vihman et al., 2004; Houston et al., 2000). By contrast, changes to the offset of disyllabic words (e.g., *bonjour* as *bongour*) affect word recognition in French- and Hebrew-learning infants, whereas changes to the onset of weak syllables do not (Hallé and de Boysson-Bardies, 1996, e.g., *bonjour* as *ponjour*;). Therefore, rhythmic cues allow infants to direct their attention toward salient portions of speech - signaled through different prominence and grouping patterning across languages - thus sustaining early lexical acquisition (Russo et al., 2021; Calignano et al., 2021a).

Lastly, rhythm also facilitates the processing of higher levels of the linguistic hierarchy. For instance, languages vary for their pattern of prominence at the phonological phrase level as well and this mainly relates to word order. Specifically, subject-verb-object languages (e.g., English and Italian; SVO) show final prominence whereas subject-object-verb languages (e.g., Turkish; SOV) show the opposite pattern. Therefore, in SVO languages, a regular sentence like ’eat apples’ is characterized by stress on the final object; by contrast, in SOV languages, a regular sentence like ’elma ye’ (*‘apple-eat’*) is characterized by stress on the initial object (Nespor et al., 2008; Langus et al., 2017). At this level of the linguistic hierarchy, rhythm emerges from the alternation of prominence between words carrying phonological phrase (PHPH) stress. Infants rely on this cue to bootstrap into the syntactic structure of their language, being sensitive to PHPH rhythmic patterning from their first months (Christophe et al., 1997) and relying on such cues to attune to their native language patterns (Bernard and Gervain, 2012).

Specifically, they rely on the way in which prominence is instantiated at the PHPH level, with French-learning infants preferring phrasal rhythm varying in duration and bilingual infants categorizing sentences based on phrasal rhythm varying in both duration and pitch/intensity according to their native languages (Bernard and Gervain, 2012; Gervain et al., 2012). Therefore, infants rely on rhythm at the PHPH level and combine this cue with other linguistic features (e.g., word frequency) to build the first representations of their native phrasal structure, bootstrapping into syntax and grammar.

In summary, linguistic rhythm is available to human fetuses since their first experience with language. At birth, signs of such early experiences can be found in their preference for familiar rhythms in language. Growing into infancy, early sensitivity to familiar rhythms guides perceptual and attentional processes toward the speech stream, at various levels of the prosodic hierarchy from metrical feet to phonological phrases. Therefore, linguistic rhythm is found to bootstrap early stages of language acquisition. Interestingly, many parallels can be made between the hierarchical structure of language and music.

Accordingly, a substantial overlap between the basic perceptual and cognitive mechanisms processing both music and speech signals have been found (Fiveash et al., 2021, outlined in the PRIMS framework, see Section **??**), consistently with the phylogenetic evolution of common neural mechanisms underlying the rhythmic processing of both signals (Kotz et al., 2018, see Chapter 2.1.1). Therefore, this led to the idea that general rhythmic abilities might be in place since early life serving the processing of complex temporal signals including music and speech (Fiveash et al., 2021; Nayak et al., 2022). Furthermore, shared underlying difficulties in general rhythm processing have been found across different speech/language disorders, leading to the hypothesis that atypical rhythmic abilities might represent an early marker of atypical developmental trajectories (Ladányi et al., 2020; Fiveash et al., 2021; Lense et al., 2021). Given the importance of identifying and targeting early underlying mechanisms of at-risk speech/language development, the link between non-linguistic rhythmic abilities in typical and atypical developmental trajectories will be outlined in the next section of this Chapter.

3.2 Developmental trajectories

3.2.1 General rhythmic abilities

As described in the previous section, music and speech rhythms share several commonalities regarding the alternation of prominence at multiple hierarchical levels (Figure 3.1). Considering speech as one of many rhythmic signals, the general definition of rhythm can be traced back to 'the product of prominence and patterning' given in the first paragraph of this Chapter (Arvaniti, 2009, p.61). Based on this definition, non-linguistic rhythms (varying in complexity from the regular pulse of a metronome to the orchestral features of musical pieces) are governed by regular pulses or beats. Beats alternate in prominence between strong and weak ones, with different possible numbers of weak beats being nested into strong, regular ones and this gives rise to the meter (Figure 3.1). Therefore, *beat* is defined as the prominent periodicity of rhythmic signals while 3.2 is the temporal structure arising from the perceptual grouping of beats in time (McAuley, 2010). Specifically, that is the structure of a perceived hierarchy of patterns of strong and weak beats alternating in time (Fitch, 2013, as syllables in speech so notes in music).

Crucially, the perceptual inference of a beat/metrical structure from rhythmic signals (Figure 3.2) requires the combined activity of several cognitive processes, including the estimation of time and duration intervals; general cognitive encoding and processing mechanisms such as working memory and attention; and movement, with the involvement of (pre)motor areas allowing for motor entrain to a beat, a peculiar ability of high vocal learner species (Grahn and Brett, 2007; Fiveash et al., 2021; Patel, 2021, see Chapter 2).

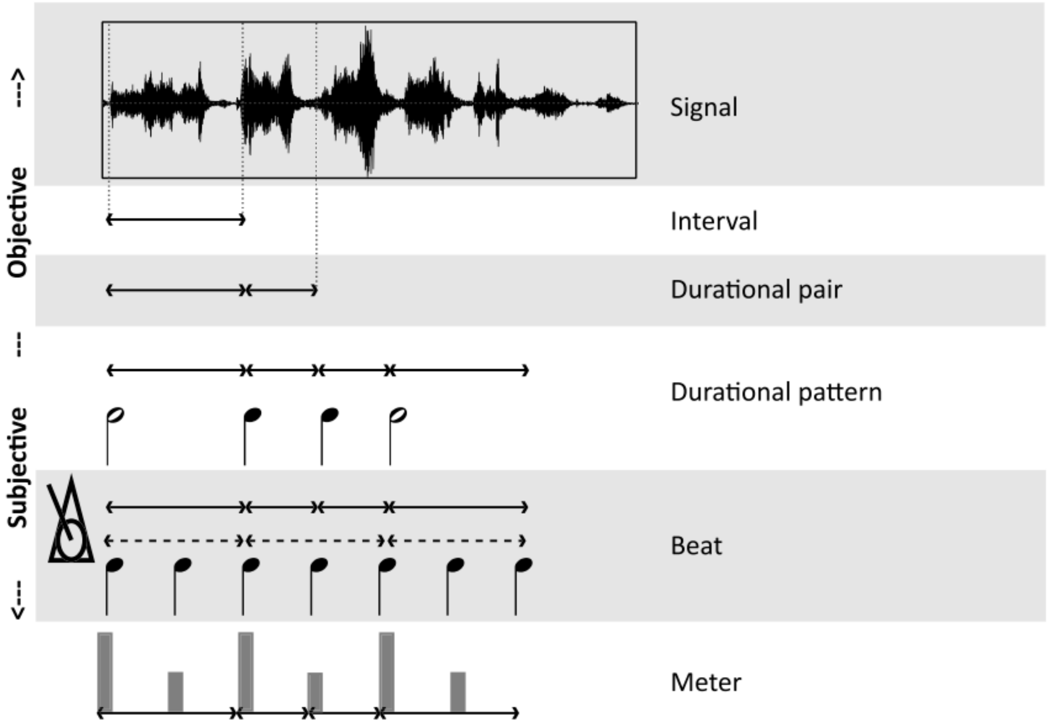


FIGURE 3.2: Objective (i.e., physical) components of the signal (top) include the Signal, the Interval between events, and the Durational pair of sequential events. From these features, the Durational pattern is perceived and then the Beat and Meter patterns are inferred (bottom). From Kotz et al. (2018).

Therefore, the processing of non-linguistic and linguistic rhythms involves domain-general abilities (including beat and meter extraction, synchronization, and motor entrainment) sustaining the temporal organization of signals unfolding in the time dimension, such as speech, music, or motor patterns. This domain-general set of abilities, defined from now on as **??**, has been found to serve many high-order functions including language, music, dance, and interpersonal interaction (Miendlarzewska and Trost, 2014; Russo and Valenza, 2021).

As Table 3.2.1 shows, many studies brought evidence in support of a link between general rhythmic abilities and language skills across different stages of development. Specifically, Table 3.2.1 summarizes forty-two studies resulting from keywords search (i.e., rhythm; language; infants; children) in scientific databases (i.e., Scopus, Google Scholar, and Pubmed) that were carried out in the past two decades (2000-2022). These studies explore related rhythmic and language abilities in typically developing individuals tested as experimental or control group. General rhythmic abilities were tested as: beat detection, motor tapping and clapping, same/different meter or rhythm discrimination, neural tracking, rise time discrimination, rhythm imitation, and rhythm production; whereas language skills were tested as: phonological awareness, reading accuracy and fluency, expressive and receptive vocabulary, grammar comprehension and production, naming or non-word repetition, perception and neural encoding of speech, morphosyntactic processing and production, and language articulation. Starting from the reported findings, several conclusions can be made.

Overall, there is an ample evidence in favor of a link between general (i.e., non-linguistic) rhythmic abilities and language skills, as shown by the large corpus of studies (N = 42) reporting signs of such relationship along development. Specifically, rhythmic abilities have been found to be related to basic language skills in young children (Kalashnikova et al., 2019, 2021; Cantiani et al., 2019; Politimou et al., 2019, e.g., phonological awareness and vocabulary at 1-4 years) and to higher literacy skills in children and adolescents (Sun et al., 2022; Lee et al., 2020; Persici et al., 2019; Goswami et al., 2013, e.g., reading accuracy, grammar processing, and morphosyntactic knowledge at 9-17 years). Furthermore, rhythm was found to be related to neural measures of language processing (Woodruff Carr et al., 2014, e.g., neural speech encoding). Therefore, rhythm seems to be related to a heterogeneous set of language skills spanning from basic to higher functions and from early to late development.

TABLE 3.1: Literature review on rhythm-related and language-related abilities

Reference	Year	age	Rhythmic measure	Language measure
Sun et al.	2022	9-11 years	behavioral and neural beat sensitivity	Phonological awareness and reading accuracy
Kertész and Honbolygó	2021	6-7 years	Sensorimotor synchronization	Phonological awareness, word reading accuracy and fluency
Kalashnikova et al.	2021	4 years	Metronome tapping	Non-word repetition, vocabulary, and letter knowledge
Eccles et al.	2021	5-7 years	Rhythm discrimination	Phonological awareness
Persici et al.	2021	6 years	Neural responses to rhythm	Expressive grammar abilities
Csaba et al.	2020	6-7 years	Metronome and un-paced tapping	Phonological awareness and reading abilities
Bonacina et al.	2020	5-7 years	Drumming to a beat and tapping rhythmic patterns	Rapid naming and phonological awareness
Swaminathan and Schellenberg	2020	6-9 years	Rhythm discrimination	Speech perception and grammar
Lee et al.	2020	7-17 years	Rhythm discrimination	Grammar-comprehension

Lê et al.	2020	8-9 years	Rhythm production	Phonological awareness, rapid automatized naming, and literacy
Cantiani et al.	2019	6 months	Oscillatory brain activity	Expressive vocabulary at 2 years
Kalashnikova et al.	2019	7-10 months	Rise time discrimination thresholds	Expressive vocabulary at 3 years
Persici et al.	2019	10 years	Tapping	Morphosyntactic processing
Politimou et al.	2019	1, 3 and 4 years	Rhythm perception and production	Phonological awareness
Ozernov-Palchik et al.	2018	5-6 years	Rhythm and metric discrimination	Phonological awareness and letter-sound knowledge
Bonacina et al.	2018	5-7 years	Synchronization abilities	Envelope encoding and literacy
Carr et al.	2017	17 years	Synchronization abilities with visual feedback	Phonological memory and reading sub-skills
Cohrdes et al.	2017	5-7 years	Sounds, short melodic or rhythmic phrases perception	Vowels, words, syllables, syntax and story perception
Vuolo et al.	2017	4-5 years	Tapping and clapping	Language processing and production

Gordon et al.	2015	6 years	Rhythm perception skills	Morpho-syntactic production
Degé et al.	2015	6 years	Rhythm and meter perception	Phonological awareness
Lee et al.	2015	9-12 years	Rhythm imitation	Phonological processing
Cumming et al.	2015	6-12 years	Beat detection and Tapping	Phonology and reading
Sallat and Jentschke	2015	5 years	Rhythmic–melodic perception	Language comprehension and production
Wieland et al.	2015	6-11 years	Rhythm discrimination	Language production
Carr et al.	2014	3-4 years	Beat synchronization	Neural encoding of speech and behavioral pre-reading skills
Chang et al.	2016	11 years	Rhythm network connectivity and behavioural rhythm discrimination	Language production
Moritz et al.	2013	5 years	Rhythm discsrimination	Phonological awareness and reading
Goswami et al.	2013	8-14 years	Beat perception	Reading

Grube et al.	2012	11 years	Isochrony deviation detection, Regularity detection and Metrical discrimination	Phonological awareness and reading abilities
Huss et al.	2011	8-13 years	Musical metrical and amplitude envelope perception	Phonological awareness and reading
Strait et al.	2011	10 years	Rhythm discrimination	Phonological awareness
Zelaznik and Goffman	2010	6-8 years	Tapping and drawing circles in time with a metronome	Language comprehension and production
Holliman et al.	2010	6 years	Rhythm discrimination	Phonological awareness and reading
Olander et al.	2010	4-6 years	Rhythmic motor timing	Language production
Corriveau and Goswami	2009	7-11 years	Paced and un-paced tapping	Vocabulary, phonological awareness, and reading
Thomson and Goswami	2008	10 years	Rhythmic discrimination, Paced and un-paced finger tapping	Phonological awareness and reading
Muneaux et al.	2008	11 years	Beat perception	Phonological processing and reading
David et al.	2007	6-11 years	Rhythm production	Reading ability

Overy et al.	2003	7-10 years	Tapping out the rhythm of a song	Spelling ability
Anvari et al.	2002	4-5 years	Rhythm discrimination	Phonological awareness and reading
Overy	2000	6-7 years	Rhythm and meter discrimination, rhythm production	Phonological skills, spelling, and reading

The general rhythmic skills related to language included a large set of abilities spanning from rhythm perception-related (Eccles et al., 2021; Anvari et al., 2002; Kalashnikova et al., 2019, e.g., beat detection, same/different judgments, rise time discrimination threshold) to rhythm production-related skills (Lê et al., 2020; Politimou et al., 2019; Lee et al., 2015; David et al., 2007, e.g., motor tapping and clapping, neural tracking, rhythm imitation, and rhythm production). Consistently with behavioral data, rhythm entrainment was found to be related to language at the neural level too (Sun et al., 2022; Persici et al., 2019; Cantiani et al., 2019; Chang et al., 2016, i.e., neural beat sensitivity, oscillatory brain activity, network connectivity). Lastly, language-relevant rhythmic skills include the motor domain as well. Indeed, sensorimotor rhythmic abilities were found to be strongly associated with language perception and production (Kertész and Honbolygó, 2021; Bonacina et al., 2020; Vuolo et al., 2017; Woodruff Carr et al., 2014; Persici et al., 2019; Cumming et al., 2015; Zelaznik and Goffman, 2010; Olander et al., 2010; Corriveau and Goswami, 2009; Thomson and Goswami, 2008; Overy, 2003, i.e., un-paced and paced tapping, clapping and drumming to a beat, tapping down the rhythm of a song, drawing circles in time with a metronome)

In conclusion, general rhythmic skills seem to be deeply related to language development. This result, emerging from the last two decades of research, is in line with the claim made by the processing rhythm in speech and music framework (Fiveash et al., 2021, i.e., PRIMS, see Chapter 2) that is, similar mechanisms underlying the processing of rhythmic signals mainly defined as precise auditory processing, oscillatory brain activity, and sensorimotor coupling. Despite the cognitive and neural dynamics underlying this cross-domain coupling is still unclear, further evidence in support of a domain-general set of rhythmic abilities serving the processing of multiple, complex rhythmic signals comes from the literature on atypical language development. Accordingly, atypical rhythmic skills have been found to be present across different neurodevelopmental disorders causing language- and motor-related difficulties (Goswami, 2011; Chang et al., 2016; Trainor et al., 2018; Carrer, 2015, e.g., developmental dyslexia, specific language disorder, autism spectrum disorder, attentional deficit/hyperactivity disorder, stuttering, developmental coordination disorder).

Unifying these findings, the Atypical Rhythm Risk Hypothesis was formulated, stating that - given the crucial role of rhythmic abilities in language development - atypical rhythm could be a risk factor for developmental speech/language difficulties (Ladányi et al., 2020, ARRH). This promising hypothesis and the evidence that contributed to framing it will be discussed in the next section of this Chapter.

3.2.2 Rhythmic abilities and atypical language development

Impairments in rhythm, timing, and synchrony/coordination have been found to occur across different neurodevelopmental disorders interfering with the acquisition of typical language abilities (Goswami, 2002; Chang et al., 2016; Trainor et al., 2018; Carrer, 2015). Specifically, through different tasks (i.e., discrimination, synchronization, and coordination) atypical rhythmic skills were linked to conditions affecting language, attention, and motor development (Lense et al., 2021, i.e., speech/language disorder, attentional-deficit/hyperactivity disorder, developmental coordination disorder, autism spectrum disorder, Williams syndrome).

Regarding the area of speech/language abilities, evidence mainly comes from the study of developmental dyslexia, developmental language disorder, and stuttering. Specifically, developmental dyslexia is a condition known to affect almost 3-10% of children with difficulties in phonological processing significantly impacting the development of typical reading abilities (Snowling, 2013, DD). A large corpus of studies, integrated into the Temporal Sampling Framework (Goswami, 2011, TSF), showed that individuals suffering from dyslexia display atypical entrainment of neural oscillations to external linguistic and non-linguistic rhythms, impacting both rhythm and phonological processing. Specifically, these individuals show deficits in synchronizing to both the speech envelope (Leong and Goswami, 2014; Goswami et al., 2010; Surányi et al., 2009) and non-speech stimuli (Cutini et al., 2016; Frey et al., 2019).

Impaired beat synchronization is also shown by individuals with dyslexia (Colling et al., 2017; Overy, 2003; Thomson and Goswami, 2008), with weaker rhythmic perception and production abilities being related to poorer phonological and reading skills (Flaugnacco et al., 2015; Forgeard et al., 2008; Goswami et al., 2013; Lee et al., 2015; Thomson and Goswami, 2008; Dellatolas et al., 2009; Muneaux et al., 2004). Furthermore, individuals with dyslexia show impaired processing of rise-time information and related inefficient neural entrainment to speech (Goswami and Bryant, 2016; Huss et al., 2011). In children, this timing deficit is particularly marked with stimuli at 2 Hz (Soltész et al., 2013) which is the frequency of accented syllable rate, crucial in language processing. Later in development, adults with dyslexia display weaker synchronization and beat perception skills (Pasquini et al., 2007; Thomson et al., 2006) with impaired low-frequency neural entrainment to speech (Molinaro et al., 2016) and non-speech stimuli (Hämäläinen et al., 2012; Lizarazu et al., 2015). Therefore, atypical rhythmic skills are widely spread among children and adults suffering from dyslexia and, specifically, a timing deficit seems to affect both rhythmic and phonological processing in this population (Ladányi et al., 2020; Goswami, 2011).

Similarly, difficulties in rhythm perception and production also affect children with developmental language disorder (Bedoin et al., 2016; Cumming et al., 2015; Sallat and Jentschke, 2015, DLD). This neurodevelopmental condition, previously identified as specific language disorder, has a prevalence rate of about 3-7% and it is characterized by poor linguistic abilities not explained by other biomedical conditions (Bishop, 2017). As for dyslexia, individuals with DLD show poor synchronization abilities when tapping to a beat (Corriveau and Goswami, 2009; Cumming et al., 2015; Vuolo et al., 2017). Moreover, the processing of low-level cues related to rhythm processing is also impaired in these children, showing atypical amplitude envelope and rise-time information processing (Corriveau and Goswami, 2009; Goswami and Bryant, 2016; Richards and Goswami, 2015). These deficits have been linked to poor performances in speech, language, and literacy measures (Corriveau and Goswami, 2009; Cumming et al., 2015; Richards and Goswami, 2015; Ladányi et al., 2020).

Similarly, children who stutter also display atypical rhythmic skills. Stuttering is a speech disorder with approximately 0.3-5.6% of prevalence in the developing population and causes repetitions or delays in sound production which disrupt the rhythmic structure of speech streams (Yairi and Ambrose, 2013). An impairment in predictive timing via sensorimotor coupling has been found as a possible underlying cause for the many rhythmic deficits displayed by these children (Hickok et al., 2011). Experimental findings sustain this hypothesis, showing: i) impaired sensorimotor coupling (Chang et al., 2016; Hickok et al., 2011), ii) inefficient timing cueing in basal ganglia circuits (Alm, 2004; Toyomura et al., 2011), and iii) behavioral difficulties in tasks related to internal time keeping (Olander et al., 2010), synchronization (Falk et al., 2015), and rhythmic discrimination (Wieland et al., 2015) in children who stutter compared to controls (Ladányi et al., 2020). Therefore, atypical rhythmic abilities are shown across different developmental speech/language disorders and seem to be related to linguistic skills via shared underlying timing processes (Goswami, 2011).

Interestingly, impairments in timing and rhythmic abilities have been found to be common also in children with speech/language impairments as comorbid deficits in other neurodevelopmental disorders affecting the area of language, attention, and motor functions. Attention-deficit / hyperactivity disorder is one of these conditions (Association et al., 2020, ADHD). With a prevalence of approximately 4-7%, this syndrome causes a pervasive pattern of inattention and hyperactivity/impulsivity. Children and adults suffering from ADHD show atypical rhythmic skills, including difficulties in perceiving, reproducing, and comparing durations (Noreika et al., 2013) and music beat perception (Puyjarinet et al., 2017). Moreover, these children show atypical synchronization in tapping to a musical beat, with underlying difficulties at the level of internal beat generation and maintenance (Noreika et al., 2013; Lense et al., 2021). Neurofunctional differences are also exhibited in the brain connectivity of ADHD phenotypes, particularly regarding connections between the basal ganglia and supplementary motor area, premotor cortex, and auditory cortex (Valera et al., 2010).

Similar findings come from individuals with developmental coordination disorder, a condition with about 5-7% of prevalence and characterized by poor coordinated gross and/or fine motor skills, significantly impairing the child's everyday functioning (Association et al., 2020). Difficulties in sensory processing of auditory, vestibular, and visual stimuli also affect these children, including relevant timing deficits such as motor and sensorimotor timing, poor predictive internal time modelling, rhythmic coordination, motor sequencing and learning, anticipation, paced tapping to visual and auditory signals (Allen and Casey, 2017; Wilson et al., 2017; Adams et al., 2017; Caçola, 2016; de Castelnau et al., 2007; Mackenzie et al., 2008; Ladányi et al., 2020). Interestingly, motor impairments have been found to be related to atypical language outcomes in children with DCD (Mirabella et al., 2017; Ladányi et al., 2020).

Evidence of atypical rhythm processing also comes from the study of individuals with autism spectrum disorder (Association et al., 2020, ASD) and, lastly, Williams syndrome (Martens et al., 2008, WS). The former is a highly heritable disease affecting about 1.5% of the population, characterized by mild to severe impairments in social functioning and communication - including possible difficulties in speech/language - as well as patterns of restricted / repetitive behaviors and pervasive sensory processing atypicalities (Association et al., 2020). WS is a rare genetic disorder affecting about 1 in 20.000 live births and it is characterized by a mixture of delay, deviance, and asynchronies in cognitive-behavioral trajectories, along with severe physical impairments; however, children with WS are found to considerably grow in language proficiency by adulthood (Martens et al., 2008; Karmiloff-Smith, 2007). Rhythmic abilities are compromised in these two neurodevelopmental disorders. Specifically, children with ASD often exhibit atypical sensorimotor coordination, atypical speech prosody, and difficulties in synchronizing with others in rhythmic activities such as rocking, clapping, or imitating movements (Franich et al., 2021; Mayall et al., 2021; Fitzpatrick et al., 2017; Marsh et al., 2013; Kaur et al., 2018; Landa et al., 2011). In children with WS, reduced rhythmic skills are often reported in tests involving same/different judgments of rhythmic patterns, musical beat extraction, and reproducing rhythmic by clapping or singing (Lense and Dykens, 2016; Martens et al., 2010; Hopyan et al., 2001; Mayall et al., 2021).

Interestingly, the rhythmic difficulties described so far are often linked to atypical language profiles across neurodevelopmental disorders. For instance, in both ASD and WS phenotype, poor synchronization skills are associated with reduced sensitivity to word stress and atypical speech prosody production, impacting turn-taking and social pragmatic (Patten et al., 2014; Nazzi et al., 2003; Klein-Tasman et al., 2009; Lense et al., 2021). Therefore, recurrent patterns of comorbidity are often shown in rhythmic and language difficulties across different neurodevelopmental disorders (Lense et al., 2021). Moreover, symptoms of one condition often appear in comorbidity with other neurodevelopmental disorders. For instance, ADHD is typically seen to occur in comorbidity with language disorder, DCD, and ASD (Piek et al., 1999; Kadesjö and Gillberg, 2001; Hawks and Constantino, 2020; Lense et al., 2021). Therefore, the simultaneous presence of similar difficulties across different conditions often appearing in comorbidity with one another has been argued to call for a transdiagnostic approach instead of a discrete, categorical diagnostic system, where neurodevelopmental disorders are investigated in terms of difficulties in general underlying mechanisms together with genetic and environmental risk factors (Mareva and Holmes, 2019; Ladányi et al., 2020). According to this view, a generally impaired ability to process rhythm and timing might be a good candidate as an early marker of language difficulties across several neurodevelopmental disorders (Ladányi et al., 2020).

In summary, different aspects of atypical rhythm processing (i.e., discrimination, synchronization, and coordination) have been identified along with speech/language impairments as co-morbid deficits in many neurodevelopmental disorders affecting the area of language, attention, and motor functions (i.e., speech/language disorder, ADHD, DCD, ASD, WS). Other than mirroring the link between rhythm and language, cognition, and motor function seen in typical development (see the previous paragraph), these findings and their interpretation following a transdiagnostic approach suggest that rhythm processing and language-related abilities likely overlap in their genetic architecture (i.e., *pleiotropy*) and neural endophenotypes, with gene-environment interactions and neuroplasticity reinforcing this link during development (Nayak et al., 2022; Ladányi et al., 2020).

Consistently, both language-related and rhythmic-related skills show moderate heritability (Andreola et al., 2021; Seesjärvi et al., 2016; Mosing et al., 2014; Drayna et al., 2001) and a large-scale genome-wide association study (GWAS) recently revealed highly polygenic architecture in the human capacity to synchronize to a beat (Niarchou et al., 2021). Specifically, this study effectively identified alleles at 69 separate loci being deferentially associated with typical versus atypical synchronization, in line with existing evidence on underlying neural mechanisms (Niarchou et al., 2021). Therefore, based on the integrated findings on typical and atypical development coming from behavioural, neural, and genetic studies, the Atypical Rhythm Risk Hypothesis posited that rhythmic abilities might vary in a continuous, normal distribution among the population and that the genetic liability for atypical rhythm might increase the risk of different developmental difficulties through genetic pleiotropy, as illustrated in the next paragraph (Ladányi et al., 2020; Lense et al., 2021; Nayak et al., 2022, see Figure 3.3).

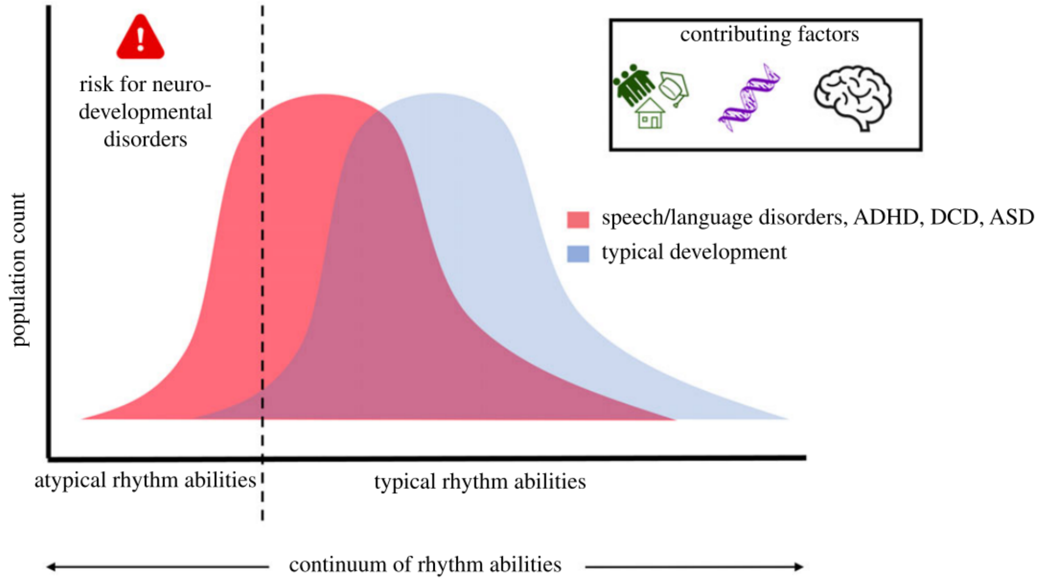


FIGURE 3.3: Rhythm (x-axis) varies from atypical to typical abilities and its link to typical (light blue) and atypical (red) development in the population. In particular, poor rhythmic skills coincide with the tail of the atypical development distribution representing a risk factor for developmental disorders. Genetic vulnerabilities, environmental factors (e.g., education, training, and social interactions) and neural processing also contribute to this interaction. From Lense et al. (2021).

The *Atypical Rhythm Risk Hypothesis*

The Atypical Rhythm Risk Hypothesis (Ladányi et al., 2020, ARRH) posits that individuals with atypical rhythm processing are at higher risk for developmental speech/language disorders. This hypothesis is supported by evidence showing that individual differences in various rhythmic tasks are related to speech and language abilities in typically developing children and adults (Lê et al., 2020; Politimou et al., 2019; Lee et al., 2015; David et al., 2007, see Section 3.2); therefore, a certain degree of overlap is likely to occur in the neural circuits underlying both functions (see Chapter 2 for the evolutionary analysis of rhythm/language networks). Consistently, similar cognitive and perceptual mechanisms support rhythm and language processing (Fiveash et al., 2021, see the PRISM framework in Chapter 2) and, lastly, rhythm impairments are shared across speech and language disorders and commonly comorbid attention and motor neurodevelopmental disorders (Goswami, 2011; Chang et al., 2016; Trainor et al., 2009; Carrer, 2015, see Paragraph 3.3). Furthermore, longitudinal studies started to find significant relationships between early rhythmic abilities and later language outcomes Cantiani et al., 2019; Kalashnikova et al., 2019; Flaugnacco et al., 2015, see Table 3.2.1).

By combining these findings, the ARRH suggests that rhythmic skills should be included in early screening programs as an index of language development. Specifically, experimental paradigms investigating a vast range of rhythmic abilities are available in the literature (Dalla Bella et al., 2017; Iversen and Patel, 2010, e.g., the BAASTA battery; the Beat Alignment Test) and can be easily administered by teachers, nurses, or paediatricians; if any atypicality is detected, this can be transmitted to health professionals for further testing. According to the ARRH, individual differences in rhythmic tests will predict significant variance in language outcomes through longitudinal designs. Moreover, it is of main interest to investigate the early stages of development to identify primary risks of atypical trajectories. Specifically, poorer rhythmic abilities are expected to predict deviant trajectories in language development and, therefore, early intervention programs involving rhythmic activities could be implemented to sustain the process of language acquisition. By embracing a family-based approach, atypical rhythmic skills in parents might be included as a risk factor together with other genetic and environmental components.

Accordingly, gene-environment interactions are of main interest for the ARRH since a partially shared genetic architecture between rhythm and language processes is expected to account for phenotypic correlations. Specifically, shared underlying genetic architecture might affect overlapping neural circuits as an endophenotype of rhythm and language abilities. In conclusion, the ARRH predicts that genetic liability for atypical rhythm increases the risk of diverse developmental speech and language problems in part through genetic pleiotropy (Ladányi et al., 2020; Lense and Dykens, 2016; Nayak et al., 2022; Fiveash et al., 2020).

Consistently with the ARRH, several findings demonstrated that short- and long-term music exposure can enhance cognitive and language performances (Schön and Tillmann, 2015; Gordon et al., 2015; Miendlarzewska and Trost, 2014; Russo and Valenza, 2021). In particular, the case of music education captured the interest of many researchers in the field of music and language cognition (Musacchia et al., 2007; Schön and Tillmann, 2015; Brown, 2012). Specifically, this line of research found increased basic auditory, cognitive, and high-functioning linguistic skills in musicians compared to individuals not receiving any professional musical education. This result led to investigating the mechanisms to transfer the effects of music on linguistic skills. The effects of exposure to musical rhythms on language processing have therefore been investigated, showing significant evidence for cross-domain transfers at multiple levels (Cason and Schön, 2012; Fiveash et al., 2021; Bedoin et al., 2016, e.g., phonological awareness, reading abilities, and grammar processing). Surprisingly, the effect of exposure to musical rhythms was also found to occur across different sensory modalities (Fotidzis et al., 2018; Liu et al., 2012; Gould et al., 2018; Schneider et al., 2011). However, these recent findings were already in the mind of early musical pedagogists who incorporated the whole body in music and rhythmic education (Juntunen and Hyvönen, 2004; Seitz, 2005; Álamos Gómez and Tejada, 2020). All this evidence and its implications for experimental and clinical practices will be discussed in the next and last section of this Chapter.

3.3 Training language with rhythm

3.3.1 Music education

Listening to a musical piece might appear as a simple and relaxing activity. However, successfully following the temporal and harmonic structure of musical signals is essential to appreciate the perceived composition. This complex process involves different perceptual, memory, and cognitive abilities and underlying neural circuits (Peretz et al., 2005). Similarly, playing music with an instrument, voice or body involves fine motor abilities, sensory feedback, and executive planning (Zatorre et al., 2007). Even processing the emotional content of music, being this perceived or produced, calls for a wide activation of brain regions implied in emotion processing and regulation (Koelsch et al., 2006; Blood and Zatorre, 2001). Therefore, perceptual, mnestic, motor, and attentive abilities together with emotional processing are involved and trained by musical experience. This has been found to produce long-term effects of musical education on different cognitive abilities, also including language (Miendlarzewska and Trost, 2014; Russo and Valenza, 2021).

First, playing music allows children and adults to develop precise auditory processing abilities. For instance, fine auditory discrimination skills of pitch, intensity, and amplitude have been found to improve in children attending musical classes (Besson et al., 2007). Moreover, these fine auditory abilities tune through exercise to the native tonality and rhythmic patterns since the first years of life (Chobert et al., 2014; Hannon and Trainor, 2007). Interestingly, fine auditory processing abilities improve as well for linguistic stimuli. For instance, syllabic discrimination skills improve in children following musical classes (Chobert et al., 2014). Therefore, musical education can train fine auditory processing skills implied in language processing (François et al., 2013; Fiveash et al., 2021, see Section **??**).

The possibility to actively practice with a musical instrument gives rise to further impacts on development. Specifically, playing an instrument as well as dancing or singing allows for training fine motor and executive planning skills, requiring accurate control over complex motor patterns over time (Zatorre et al., 2007). Furthermore, sensorimotor abilities are also impacted by musical practices: a fine multi-sensory integration is required to integrate the auditory product of playing an instrument into the motor activity generating the sound. In other words, the pattern of motor actions implied in generating a musical piece by playing an instrument or singing is continually adjusted based on the auditory information produced by the signal. That is, a fine feedback mechanism is trained by musical education and has been found to lead to functional and structural changes in sensorimotor brain circuits (Luo et al., 2012).

As described in Chapter 2, sensorimotor coupling is found to play an important role in human synchronization abilities (Patel, 2021) and language (Fiveash et al., 2021, see Chapter 2). Accordingly, sensorimotor coupling sustains the ability to synchronize motor patterns to an external beat (e.g., tapping to a metronome) involving motor, sensory, and attentive control over time (Coull, 2004). Consequently, the ability to effectively allocate attention at different time points (Astheimer and Sanders, 2009; Russo et al., 2021, i.e., Temporal Selective Attention) has been pointed out as a potential underlying mechanism in the cross-domain transfers of benefits from musical to language development. This is in line with the claim made by the Dynamic Attending Theory (Jones et al., 2002; Jones, 2018) suggesting that external rhythmic regularities allow for attuning the attentional system to the structure of the signal, facilitating temporal prediction of upcoming events by allocating temporal attention to salient expected points in time.

Similar findings were reported in the Temporal Sampling Framework (Goswami, 2011, TSF, see Section 3.2.1), highlighting the role of the endogenous entrainment of neural oscillations at multiple hierarchical levels, and in the PRISM framework, pointing out this mechanism as one of the main ones underlying language processing together with auditory processing and sensorimotor coupling (Fiveash et al., 2021, see Section 3.2.1 and Chapter 2). As these theories and evidence suggest, since both music and speech appear to be tracked in the brain via similar neural mechanisms, the sustained exercise of synchronization, sensorimotor coupling, and temporal attention in musical education is hypothesized to boost the processing of speech and language signals in young and adult musicians.

Consistently, tapping in synchrony to an external beat is found to correlate not only with attentive measures but also with reading abilities in children and adolescents at typical and atypical development (Tierney and Kraus, 2014; Repp, 2005, see Section 3.2). More in general, the benefits of musical education on language processing have been found across different levels (Besson et al., 2011). For instance, the aforementioned auditory processing is enhanced at the phonemic (Strait et al., 2013), syntactic (Jentschke and Koelsch, 2009), and grammatical (Moreno et al., 2009) levels. Moreover, verbal proficiency, lexical memory, second language acquisition, and reading abilities as well seem to be positively affected by musical practice (Schlaug, 2011) with possibly related benefits for learning and academic performance (Moreno et al., 2009). Therefore, several findings suggest musical education to be associated with improvements in different linguistic abilities spanning from auditory processing to lexical learning and reading skills. One might explain this relationship by pointing out that better linguistic abilities and, more in general, higher intelligence or cultural levels could be the cause instead of the effect of better performance in musical practice (Schellenberg, 2011). For sure, general cognitive abilities and socioeconomic status must be considered in studying the relationship between musical education and language trajectories, together with other mediating factors such as age, motivation, reinforcement, and teaching modalities (Miendlarzewska and Trost, 2014; Russo and Valenza, 2021).

However, even if this relationship followed the opposite direction (i.e., higher cognitive abilities facilitating musical education) this would not go against the claim made by developmental and evolutionary theories sustaining that shared underlying cognitive and neural mechanisms support the processing of both music and language (Goswami, 2011; Fiveash et al., 2021; Patel, 2011). In fact, it is likely the case that training general underlying mechanisms (i.e., auditory processing, neural entrainment, and sensorimotor coupling) leads to benefits in both language and music domains at the same time (Patel, 2011). Moreover, language and music also share considerable similarities in terms of signal structure, being both hierarchically organized and unfolding in the time dimension (see section one of this Chapter). Therefore, starting from the documented link between musical education and general cognitive and language abilities, researchers started to design training programs based on music exposure to treat cognitive and language neurodevelopmental disorders (Brandt et al., 2012; Overy, 2003; Srinivasan et al., 2015).

Surprisingly, the effects of musical rhythmic exposure were not limited to long-term training; indeed, even short-term exposure to rhythm has been associated with improved performance in language and language-related tasks (Cason and Schön, 2012; Cason et al., 2015b; Fiveash et al., 2020). As discussed in the next paragraph, shading light on the cross-domain effects of music and language training is not only highly informative about possible shared cognitive and neural mechanisms but also useful for clinical application, as hypothesized in the ARRH (Ladányi et al., 2020, see Section 3.3). Therefore, the cross-domain effects of music on language and the following implications for development will be discussed in the next paragraph.

3.3.2 Cross-domain and cross-sensory effects

Previous in this Chapter, rhythm has been defined as 'the product of prominence and patterning' (Arvaniti, 2009, p.61, see Section 3.1.1). Moreover, several theories proposed that: i) endogenous neural oscillations in the brain entrain to rhythmic signals, and ii) that this ability allows the cognitive system to operate predictions on upcoming events improving the processing of salient portions in the signal (Jones, 2018; Goswami, 2011, DAT, TSF. See Chapter 2). Interestingly, these theories also suggested that endogenous oscillations persist - for a short time window - after the external stimulus has stopped (Jones, 2018). This mechanism has been hypothesized to provide a neural basis for the influence of a short exposure to a rhythmic cue on subsequent signal processing, as studied in rhythmic priming paradigms (Barnes and Jones, 2000; Jones, 2018; Kösem et al., 2018). Specifically, rhythmic priming studies allow for studying cross-domain effects of short-term rhythm exposure on the subsequent processing of other signals including speech and language, in typical and atypical development (Cason and Schön, 2012; Cason et al., 2015a; Fiveash et al., 2020; Bedoin et al., 2016). Moreover, the effect of possible confounding factors, common to long-term training (see Paragraph 3.3), is limited with single, experimental sections. Given the fact that music training and priming studies might investigate slightly different mechanisms in play, a complementary approach must be taken into account in studying cross-domain effects of musical rhythm on language processing (Schön and Tillmann, 2015).

During the last decade, rhythmic priming studies have shown that the rhythm of a prime cue can influence the processing of subsequent signals (Barnes and Jones, 2000; Jones et al., 2002; Kösem and Van Wassenhove, 2017). Specifically, a rhythmic cue has been found to facilitate the processing of subsequent matching signals compared to mismatching or irregular cues (Cason and Schön, 2012; Cason et al., 2015a; Fiveash et al., 2020; Bedoin et al., 2016).

Considering linguistic signals, behavioral studies showed that a matching rhythmic prime facilitates speech processing at the phonemic (Cason and Schön, 2012; Cason et al., 2015a), syllabic (Fiveash et al., 2020; Gould et al., 2017), and syntactic level (Chern et al., 2018; Fiveash et al., 2020). Specifically, participants were faster in detecting target phonemes in non-words or sentences when rhythmic primes matched the phoneme presentation at the beat or metric rates (Cason and Schön, 2012; Cason et al., 2015a; Fiveash et al., 2020). Moreover, they were more accurate in grammar processing and faster in reading following regular vs irregular rhythmic presentations (Chern et al., 2018; Fiveash et al., 2020). Therefore, musical rhythm can influence the behavior displayed toward linguistic information. Similar evidence comes from electrophysiological studies (Cason and Schön, 2012; Canette et al., 2020; Zhang and Zhang, 2019). For instance, EEG studies showed that the P300 and N100 components are enhanced when phonemes are presented off- vs on-beat, suggesting a larger violation detection for off-beat phonemes (Cason and Schön, 2012). Moreover, studies on time-frequency analysis confirmed the role of neural oscillations behind the observed effects by showing enhanced phase-locking of neural oscillations at the frequencies of speech stimuli after regular cues (Lanzilotti et al., 2019; Gordon et al., 2011). Therefore, short-term exposure to musical rhythm can boost the processing of linguistic stimuli at the behavioural and cortical levels through neural entrainment.

Interestingly, rhythmic priming has been found to facilitate language processing also in at-risk and atypical development. Specifically, musical primes improved the syntactic, grammatical, and phonological processing as well as the sequencing ability and sentence repetition of speech in congenitally deaf children, a population at particular risk for syntactic perception and production deficits (Bedoin et al., 2016; Cason et al., 2015a). Rhythmic musical primes also improve grammatical processing in children and adults with dyslexia and DLD both at the behavioural and neural level (Przybylski et al., 2013; Ladányi et al., 2021; Bedoin et al., 2016; Canette et al., 2020, 2019). Children who stutter as well were found to benefit from external rhythmic cues in producing speech and singing (Toyomura et al., 2011; Falk and Dalla Bella, 2016; Wan et al., 2010). Therefore, rhythmic priming paradigms offer promising results in treating language-related difficulties in neurodevelopmental disorders as well as in at-risk and typically developing populations.

Surprisingly, the effect of a rhythmic prime has been found to occur not only across domains (i.e., from music to language) but also across sensory modalities. Specifically, studies reported cross-modal effects of auditory rhythm on visual word recognition (Fotidzis, 2020; Bolger et al., 2013; Brochard et al., 2013). Moreover, visual rhythmic cues have been found to aid subsequent auditory processing (Ten Oever et al., 2014; Liu et al., 2012; Gould et al., 2018).

Interestingly, multimodal rhythmic stimulation was found to enhance rhythm perception, with audio-motor rhythmic priming enhancing subsequent language processing compared to auditory only priming (Geiser et al., 2010; Manning and Schutz, 2013; Su and Pöppel, 2012; Cason et al., 2015a). As shown in Chapter 2, sensorimotor abilities have been found to play a crucial role in rhythmic processing (Fiveash et al., 2021). Therefore, tactile and motor abilities might represent a useful channel to transfer the effects of rhythmic perception to language via audio-motor priming. Accordingly, a motor rhythmic prime was shown to enhance following linguistic processes such as lexical-decision and speech production abilities in children (Rueschemeyer et al., 2010; Cason et al., 2015a). Moreover, neurophysiological evidence showed the involvement of haptic and motor systems in predictive time functions, facilitating neural and behavioral processing in both musical and linguistic contexts (Morillon and Baillet, 2017; Schneider et al., 2011; Tierney and Kraus, 2014). Interestingly, this cross-modal dynamic (i.e., sensorimotor to auditory) also goes in the opposite direction (i.e., auditory to sensorimotor), with auditory melodic priming enhancing subsequent motor sequencing performances including speech production (Stephan et al., 2016; Cason et al., 2015b). Therefore, these results show: i) evidence for a key role of the sensorimotor system in providing temporal information to sensory regions, thus enhancing perceptual and behavioural processing; and ii) that an internal representation of perceived rhythmic patterns seems to drive the following processing of stimuli regardless of the encoding sensory modality (i.e., cross-sensory dynamics).

In the previous Chapter, the roots of human rhythmic abilities were traced back to the prenatal experience with vestibular-tactile-somatosensory (VTS) stimuli in the womb (Provasi et al., 2014, see Section 2.3.1). Moreover, the effect of bouncing and touching infants on early rhythmic and language development were also discussed in the same Section (Phillips-Silver and Trainor, 2005; Rocha et al., 2021; Seidl et al., 2015, see Paragraph 2.3.2). In this Chapter, VTS and sensorimotor abilities were shown to be related to language skills across typical (Kalashnikova et al., 2021; Kertész and Honbolygó, 2021) and atypical development (Persici et al., 2019; Corriveau and Goswami, 2009). Moreover, it has been shown that musical training involving both long- and short-exposure to sensorimotor rhythms enhances processing and production abilities in musical as well as linguistic contexts (Cogo-Moreira et al., 2013; Herrera et al., 2011; Cason et al., 2015a,b). Therefore, - across Chapters - it might become clear the main claim of this work; that is, rhythmic abilities and VTS/sensorimotor experiences might be foundational to the language from the earliest stage of acquisition and throughout both the following developmental processes. The general idea that rhythm and movement are deeply connected is not new. Despite being relatively under-explored in developmental sciences (Bremner and Spence, 2017, see Section 2.3 and Figure 2.5), it is in fact common across anthropological theories and pedagogical techniques, offering the interesting insights illustrated in the next and last Section of this Chapter.

3.3.3 Developing rhythmic skills through the body

Among the brilliant personalities of his time, Aristotle was probably the first one to assert that 'motion follows the sound' (Helmholtz, 1954; Phillips-Silver and Trainor, 2007). Accordingly, humans are known to engage in shared musical and rhythmic behaviours in the forms of dance or synchronized movements since ancient times and across almost all cultures around the world (Kotz et al., 2018, see Chapter 2). Nowadays, we often listen to music experiencing an inner tendency to instinctively tapping the feet, whistling, singing, or dancing (Wallin et al., 2001; Phillips-Silver and Trainor, 2007). Interestingly, these forms of shared rhythmic behaviours were hypothesized to prompt the first forms of communication, resulting in the neural networks underlying both rhythm and language processing by means of gene-culture adaptation (Kotz et al., 2018; Patel, 2006, 2021, see Chapter 2). An extended discussion on the evolutionary theories linking shared rhythmic behaviours to language abilities in high vocal learners has been reported in the first Chapter of this dissertation (see Section 2.1.1). However, a cross-reference to these lines of research is made at this point to support the claim - crucial to this work - that nowadays advances in developmental cognitive neuroscience should always consider the most general principles governing nature and humans when focusing on specific, sometimes microscopic phenomena.

Specifically, even though its peculiar nature as a linguistic signal, human speech is at the same time a complex and rich sound varying across the time dimension, with pitch, harmonic, and rhythmic features comparable to many other musical signals as well as further animal linguistic systems (Brandt et al., 2012; Doupe and Kuhl, 1999; Petkov and Jarvis, 2012; Santolin et al., 2019). Accordingly, the first approaches with speech occur in terms of prosodic features perceived in the maternal womb (e.g., rhythm and low frequency of speech) and, after birth, mainly in terms of sounds. Only later in development, infants start to crack the speech code finally mapping sounds into meaning (Kuhl, 2004). Within this framework, all the aforementioned commonalities in signal structure, cognitive processing, and neural underpinning of music and language should appear as logical. Furthermore, the mind–body duality proposed by R. Descartes (1985) has been criticized for decades (Damasio, 1994, 2001); however, the idea that the human mind always comes with a body even in the finest cognitive processes still surprises.

Talking about language, it is extremely rare to include the body into discussions. However, the heterogeneous literature on musical rhythm and language development described so far should have brought a clear example of the deep connection existing between rhythm, language, and the body. Accordingly, body perception and movements can be seen as a powerful channel for learning and enhancing music, rhythm, and language abilities. This has been shown also by the cross-domain and cross-sensory training practices involving long-term music classes as well as short-term music exposure through priming paradigms (Cogo-Moreira et al., 2013; Herrera et al., 2011; Cason et al., 2015a,b, see Paragraph 3.3.2). Surprisingly, these exciting findings were intuitively put into practice by music pedagogists in the early 1900. Among other pioneers, Jaques-Dalcroze (1865-1950) was the first author in music pedagogy who considered listening as a process involving the whole body. Specifically, he studied the origins of sound perception in the kinesthetic process of receiving air vibrations in the eardrum as well as in the rest of the body through mechanoreceptors (Stubley, 1999). Based on this notion, he developed a teaching style where body movements and sensations were crucial to developing rhythmicity in children and kids. He framed his body of knowledge as Dalcroze Eurhythmics and influenced the following educational strategies of his time (including the work of Z. Kodály, C. Orff, and E. E. Gordon). Nowadays models of embodied learning and cognition still echo the basic principles of the Dalcroze Eurhythmics, highlighting the role of the body in making sense of the world through development (Juntunen and Hyvönen, 2014; Álamos Gómez and Tejada, 2020; Russo and Valenza, 2021).

Therefore, rhythm processing is fundamental to language learning in typical and atypical development. Moreover, music rhythmic training can contribute in treating speech/language difficulties as well as other neurodevelopmental disorders involving the areas of attention, motion, and language. This led to considering atypical rhythm processing as a potential risk factor for later language disorders.

Short-term exposures to rhythm were also found to extend the cross-domain effects of rhythm on following language processing, with interesting evidence for cross-sensory transfers. Therefore, training rhythmic skills across modalities seems to boost language learning. Despite its central role in rhythm perception, VTS experiences and sensorimotor abilities are still less investigated compared to other sensory modalities. Yet, the strong prompt of the body experience in perceiving rhythm has been largely included in early music pedagogy with even more zeal than nowadays techniques. This is surprising considering that VTS experiences and sensorimotor abilities are natural channels for infants and children in investigating the world, given that they are highly familiar since the earliest experiences in the womb. Furthermore, VTS experiences and sensorimotor abilities can be crucial for infants and children with sensory deprivations (e.g., early blind or deaf individuals) as well as for people struggling with sensory processing (i.e., individuals with ASD). For these and many other reasons illustrated throughout this dissertation, VTS experiences and sensorimotor abilities in processing rhythm are worthy of further investigation. This is the prompt that motivated the present work. Therefore, the second part of this dissertation will illustrate the experimental research conducted to investigate VTS and sensorimotor rhythmic skills as a possible tool for sustaining language processing throughout development.

Part II

EXPERIMENTAL SECTION

Introduction to the Experimental Section

In the Theoretical section of this work, rhythmic abilities have been explored from an evolutionary perspective, shedding light on the common neural and cognitive mechanisms underlying both rhythmic and linguistic processing in humans (Chapter 2). Then, the ontogenetic evolution of rhythmic skills has been traced back to the earliest experiences with vestibular-tactile-somatosensory (VTS) rhythms perceived in the womb (Chapter 3). Moreover, the effects of processing early available sound features of human speech were investigated, highlighting the role of rhythmic structures in bootstrapping language development. Then, atypical rhythmic abilities were identified among neurodevelopmental disorders involving language difficulties as a primary deficit or in comorbidity with other attentional and motor difficulties. Accordingly, the benefit of training general rhythmic abilities on language development was discussed, both considering long-term training and short-term exposure to rhythms across sensory modalities. Lastly, the involvement of the body through VTS and sensorimotor abilities was highlighted in the last section, as intuited by last century musical pedagogists and confirmed by the growing corpus of evidence on early rhythmic, somatosensory, and sensorimotor processing. Nevertheless, a restricted corpus of studies specifically focused on VTS and sensorimotor abilities in rhythm and language development. Therefore, the present work aimed at investigating the role of the body through VTS and sensorimotor rhythmic abilities across language development. Specifically, VTS rhythmic abilities will be firstly assessed through a custom made, vibrotactile tool for music perception specifically designed for infants and toddlers based on the recent advances in the field of human-computer interaction, as illustrated in Study one (Chapter 4). In Study two, early linguistic abilities will be investigated in the same cohort of infants, specifically testing the role of phonological and prosodic features of speech (Chapter 4).

Then, the link between rhythmic and linguistic abilities will be explored between the two tasks, to investigate whether the former might be informative about the latter. In the third and last Study of this Chapter, cross-domain (i.e, from music to language) benefits resulting from rhythmic exposure across sensory modalities (i.e., from VTS to auditory signals) will be tested through a rhythmic priming paradigm (Chapter 4). Therefore, Chapter 4 will investigate VTS rhythmic abilities and their link as well as their effect on the emerging linguistic abilities of infants and toddlers. In the following Chapter, the effect of sensorimotor rhythmic abilities on complex linguistic behaviors including reading and writing skills will be examined in young adults (Chapter 5). Specifically, the effect of sensorimotor synchronization to an external beat will be evaluated while administering two screening tasks for developmental dyslexia (lexical decision and dictation under phonological suppression; Study four and Study five). Therefore, Chapter 5 will provide insights into the role of sensorimotor rhythmic abilities in literacy skills with possible implications for early screening and training practices for neurodevelopmental disorders. In conclusion, Chapter 4 and Chapter 5 together will bring evidence on the role of VTS experiences and sensorimotor abilities across language development.

Chapter 4

Linking VTS rhythm and linguistic abilities in infancy

As stated in the previous Chapters, first experiences with rhythm begin already in the womb, where fetuses are exposed to different sources of stimulation, including vestibular, tactile, and somatosensory (VTS) input produced by the maternal body (Lecanuet and Schaal, 2002; Provasi et al., 2014). The VTS experience includes tactile input perceived through cutaneous sensation, interoception, and proprioception (Gallace and Spence, 2014). Interoception provides sensations (typically unconscious) about organ functioning, homeostasis, digestion, and respiration (Craig, 2009); proprioception informs about how the body and limbs are placed or moving in space, shaping body and limbs' representation; cutaneous stimulation is integrated with interoception and proprioception to interpret our tactile environment through haptics (Bremner and Spence, 2017). All these signals are processed in the primary somatosensory cortex, giving rise to somatosensory experience (Penfield and Rasmussen, 1950). Further sense of body position, velocity, and duration of motion comes from vestibular input, which originates in the labyrinth, are refined in brainstem and cerebellar circuits, and then are relayed to the cortex (Choi et al., 2021). Together, these components are defined as vestibular-tactile-somatosensory (VTS) experience and are here hypothesized to play a role in infant rhythm perception (see Chapter 2). Specifically, this hypothesis is based on findings showing that the VTS system is the first to develop in the human fetus (Humphrey, 1964; Bremner and Spence, 2017) and, after the third trimester of gestation, it contributes to the perception of rhythmic signals from intrauterine and extra-uterine environments through bone conduction (Sohmer et al., 2001; Granier-Deferre et al., 2011).

Therefore, ‘hearing’ occurs through VTS mechanisms and early rhythmic experiences are mainly provided through this sensory modality (see Section 2.3.1). Accordingly, near term fetuses can differentiate between VTS rhythms generated by the maternal body movement (Cito et al., 2005) and premature infants adapt their breathing to the rate and acceleration of VTS stimuli (Zimmerman and Barlow, 2012). Moreover, caregivers from all cultures use VTS rhythms to calm newborns, for instance by moving them back and forth or by rocking them while singing or walking (Provasi et al., 2014). A recent study demonstrated that the cadence of parent movements and walking is related to the spontaneous motor tempo of young infants, demonstrating that the exposure to passive rhythmic actions (experienced both prior and after birth) shapes first rhythmic behaviors in infancy (Rocha et al., 2021). Prior to this, the seminal work from Phillips-Silver and Trainor (2005) showed that 7-months-old infants who were bounced in synchrony to a given rhythm preferred to listen to an auditory stimulus matching that rhythm compared to a non-matching one. Therefore, the authors concluded that passive movements generating VTS stimulation influence the perception of auditory rhythm in infants (Phillips-Silver and Trainor, 2005). After this work, the research investigating VTS rhythmic stimulation kept growing with further evidence coming from adult participants leading to the nowadays dynamical, embodied, and ecological approaches to rhythm processing (Tichko et al., 2021, see Section 2.3.2).

Interestingly, the core involvement of body movement and perception is also in line with the recent neurocognitive models claiming shared cognitive and neural underpinning of rhythm and speech perception (Fiveash et al., 2021; Ladányi et al., 2021, e.g., the PRISM framework, the ARRH). Specifically, these models suggest that, together with auditory processing and neural entrainment, sensorimotor experiences are crucial to the processing of rhythmic structure in both musical and linguistic signals (Fiveash et al., 2021; Ladányi et al., 2020, see Chapter 2).

Consistently, sensorimotor rhythmic impairments have been found to be shared among neurodevelopmental disorders impacting typical language acquisition (Ladányi et al., 2020; Goswami, 2011; Lense and Dykens, 2016, i.e., speech/language disorder, ADHD, DCD, ASD, WS; see Chapter 3). Furthermore, rhythmic training involving sensorimotor experiences were found to be useful in treating language-related skills across neurodevelopmental disorders (Neves et al., 2022; Gordon et al., 2015, see Chapter 3). Similar results come from typically developing infants, whom rhythmic abilities relate to different language skills through development (see Table 3.2.1 in Chapter 3).

Accordingly, improvements in language-related skills were found in typically developing children attending music classes, music training, and even after short-term exposure to rhythmic patterns across modalities (Miendlarzewska and Trost, 2014; Russo and Valenza, 2021; Cogo-Moreira et al., 2013; Herrera et al., 2011; Cason et al., 2015a,b). Crucially, the possibility to benefit from rhythmic cues and a more general sensitivity to rhythm across modalities has been hypothesized to support language development with atypical rhythm processing possibly targeting atypical or at-risk developmental trajectories (Ladányi et al., 2020; Goswami, 2011; Lense and Dykens, 2016). This might be explained not only by the cited shared cognitive and neural mechanisms underlying rhythm and language processing but also by the overlapping genetic architecture (i.e., pleiotropy) rhythm processing and language-related abilities likely share (Nayak et al., 2022; Ladányi et al., 2020). This would be in line with the moderate heritability of language-related and rhythmic-related skills (Andreola et al., 2021; Seesjärvi et al., 2016; Drayna et al., 2001) and with large-scale genome-wide association studies (Niarchou et al., 2021, GWAS, see Chapter 3). Therefore, the recently formulated atypical rhythm risk hypothesis (Ladányi et al., 2020, ARRH, see Chapter 3) suggests that early rhythmic skills might be included in screening and training programs as possible identifying and targeted markers of development trajectories (Ladányi et al., 2020; Fiveash et al., 2020).

However, further research is needed on the predictive strength of early rhythmic skills in language development as well as on the large individual variability often encountered in infant performances. Moreover, VTS rhythmic abilities are still less under-explored compared to other sensory modalities despite their core involvement in rhythm – and possibly in language – development. Experimentally investigating VTS rhythmic abilities would thus lead to better understand the commonalities between rhythm and language processing with further insights for early screening and training practices for language development in infancy and early childhood.

Therefore, the present Chapter will investigate this topic by means of three different studies with typically developing infants and toddlers. In the first Study, VTS rhythmic abilities will be evaluated by means of a vibrotactile system for music perception specifically designed for infant participants. This system was made based on an extended literature review on the nowadays techniques of sensory substitution and human computer interaction in the field of music technology, as described in Study one. Study two will investigate linguistic abilities (i.e., phonological and prosodic perception) in the same cohort of infants, relating individual performances between the two tested abilities (i.e., VTS rhythm and language processing). Lastly, Study three will examine the cross-domain (i.e., musical rhythm to language) effect of a rhythmic priming across sensory modalities (i.e., VTS to auditory input).

4.1 Study 1: VTS rhythmic abilities in infancy

In this Study, rhythmic abilities in vestibular-tactile-somatosensory modality will be evaluated in infants and toddlers by means of a vibrotactile system for music perception and a metric discrimination paradigm. Specifically, infant attentive behaviours displayed toward rhythmic stimuli will be collected as a measure of encoding and discrimination and derived from behavioural (i.e., looking times) and physiological (i.e., pupillometry) measures collected through the eye-tracker. Indeed, this Study brings together the recent advances from the music technology and sensory substitution field in adult perception and cognition with the finest techniques of the infant research field, allowing for studying perceptual and attentive processing in preverbal infants. This Study was motivated by an extended literature review made in the two scientific fields according to the interdisciplinary nature of my PhD program in Brain, Mind and Computer Science at the University of Padua. In developing the vibrotactile system for music perception as well as in testing infant and toddler participants, I have been supervised by researchers from the developmental Cognitive Neuroscience and the Computational Sonology fields (i.e., Prof. Eloisa Valenza, Prof. Antonio Rodà, Prof. Barbara Arfé, Dr Giulia Calignano, Dr Filippo Carnovalini). Facing the challenging time of COVID-19 pandemic, this Study – as the following two with infant participants – was tested thanks to the contribution of 5 kindergartens as part of the Cooperative Now Project. Moreover, useful insights come from inspiring discussions with researchers from abroad during research stays, conferences, and project presentations.

4.1.1 Aim and research questions

The aim of this Study was to test VTS rhythmic abilities in infancy. Specifically, it investigated whether infants are able to discriminate between different meters only based on VTS input and how this ability changes across development (from 7 to 35 months of age). Given the central role of VTS input in the perinatal period, infants were expected to be able to tease apart VTS rhythms with novel versus recently familiarized underlying meters. According to the literature on infant research, differences are expected in terms of behavioral (i.e., looking times) and physiological (i.e., pupillometry) measures displayed toward novel vs familiar stimuli in the test phase of the paradigm. Rhythmic abilities in other sensory modalities including seeing and hearing point toward positive expectations from performances with VTS stimuli (Hannon and Johnson, 2005; Brandon and Saffran, 2011). Moreover, VTS inputs were found to shape infant rhythmic behaviors in early stages of development (Phillips-Silver and Trainor, 2005; Rocha et al., 2021). Therefore, even though no study to our knowledge has investigated VTS rhythmic processing alone, evidence of efficient VTS rhythmic abilities are here expected to be found, bringing new knowledge on the perceptual and cognitive abilities of young infants. Accordingly, differentiating among different meters in rhythm is essentially a cognitive process that implies the extraction of hierarchical patterns of strong and weak beats alternating in time (Fitch, 2013, see Section 3.2.1). This cognitive ability is crucial in processing rhythmic signals such as music and speech (Kotz et al., 2018). Therefore, testing this cognitive ability across sensory modalities will help to shed light on the extent to which rhythmic abilities might be considered as a set of general, cognitive and perceptual skills serving the processing of a vast range of signals (including music and language). Lastly, given the considerable developmental changes occurring across infancy, VTS rhythmic abilities are expected to change considering age as a continuous predictor. In the interim discussions of this Study, the nature and direction of these changes will therefore be debated in terms of maturation and enculturation, two crucial processes in the gene-environment interactions shaping possible developmental trajectories.

4.1.2 Method

Participants

Forty-five infants (21 females, M age = 661.6 days, SD = 192.6) were recruited from kindergartens from Padua (IT) and adjacent municipalities as part of the Cooperative Now Project (www.coopnow.it). The inclusion criteria for all participants were: to be in good health, to have no sensory or neurological disorders or any familiar language disorder, and to be native Italian speakers. Eight participants were excluded from the analysis because of fussiness/crying (n=5) or because data did not meet the inclusion criteria set prior to data collection of at least 1 test trial per condition (n=3). The final sample included 37 participants. The research protocol was approved by the Ethics Committee of the Human Inspired Technology research centre, University of Padua (protocol number: 2020_66R1). The research was conducted in accordance with the principles of the Declaration of Helsinki. Caregivers signed a consent form before taking part in the experiment while verbal consent was asked to participants before starting the experimental section.

Stimuli

Eight different musical rhythms were originally synthesized as audio tracks with the library SoCal drum sound in GarageBand. The underlying meter of the rhythms could be a quadruple or a triple meter (Hannon and Johnson, 2005). The quadruple meter is a version of a double meter characterized by four primary beats (**1**234 - **1**234 - **1**234 - **1**234). By contrast, the triple meter is characterized by three primary beats (**1**23 - **1**23 - **1**23 - **1**23). Four different rhythms were generated for each meter, varying in the distribution of events and accents across the units: for rhythms characterized by a quadruple meter, events and accents occurred more frequently every two or four units; whereas, for rhythms characterized by a triple meter, events and accents occurred more frequently every three units (Hannon and Johnson, 2005).

Each rhythm had a duration of 6 seconds and was displayed as a VTS stimulus by the vibrotactile system within a single trial together with a randomized image presented on a monitor. Images consisted of ten, static and coloured cartoon figures presented on a neutral background to attract the visual attention of the infants and allowing the eye-tracker to collect the data. The image areas corresponded precisely to the areas of interest (AOI), which measured 10 × 10 cm (9.554 deg) and remained visible throughout the trial. The visual stimuli were equated for luminance to reduce luminance-induced variability in the pupillometry measurements (Mathôt and Vilotijević, 2022, Figure 4.1).

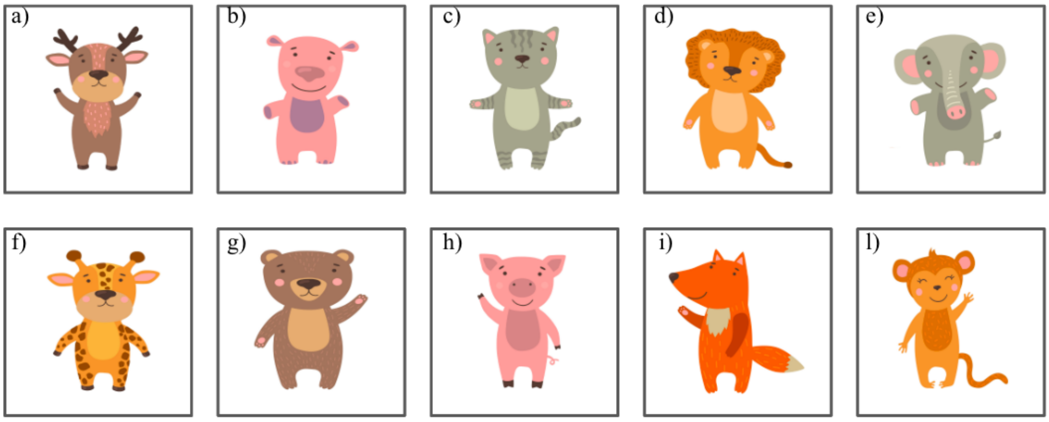


FIGURE 4.1: Visual stimuli from Study 1. Top: ten (from *a* to *l*) static, cartoon figures presented on a neutral background to attract infant visual attention. From: (freepik.com). Stimuli from *a* to *f* were presented in the familiarization phase of the paradigm, while stimuli from *g* to *l* in the test phase.

Apparatus

The experiment was programmed and presented through Open Sesame software 3.1 (Mathôt et al., 2012). Visual stimuli were displayed on a 27-inch monitor (Philips 300 × 300). Rhythms were presented in VTS modality through the Infant vibrotactile system, a music transduction device specifically developed for this Study, inspired by the Model Human Cochlea (MHC) from Karam et al. (2009b). This system is a sensory substitution technique for music perception that translates auditory information into vibrotactile stimuli using a tactile display. Music is displayed through voice coils facilitating the direct translation of auditory information onto multiple discrete channels projected toward different portions of the back (Karam et al., 2009b).

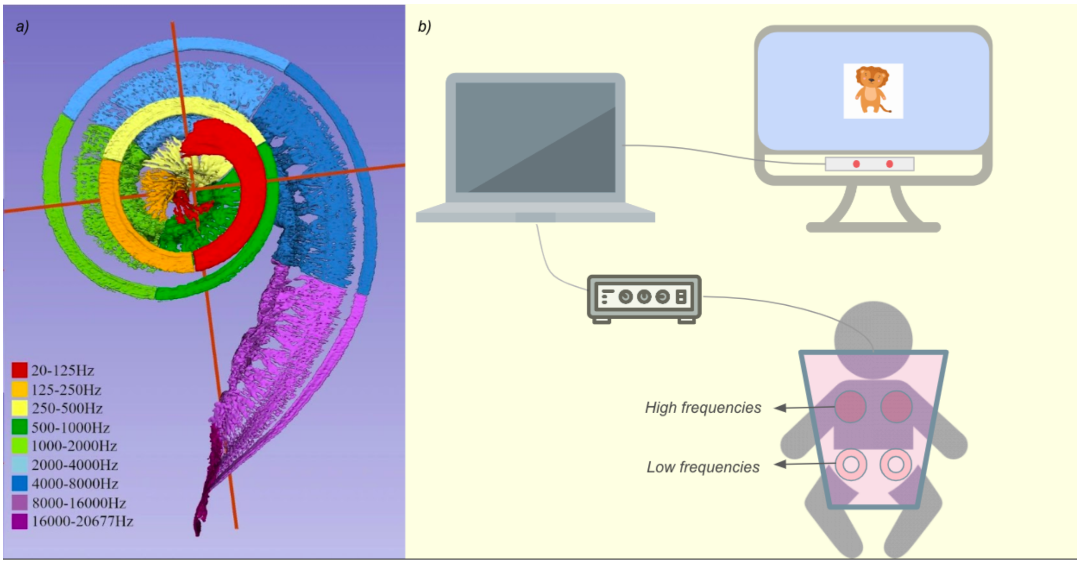


FIGURE 4.2: The Infant vibrotactile system and its implementation in the experimental set-up. The frequency map of the basilar membrane (*a*, adapted from Li et al. (2021)) inspired the topographic projection of the signal (Karam et al., 2009b). The experimental set-up *b* included: the laptop running the experiment, the monitor displaying the visual stimuli, the eye-tracker collecting data, the amplifier, and the Infant Vibrotactile System transmitting VTS stimuli in low and high frequencies through actuator channels to the upper and low part of the infant's seat-back.

The MHC system has been successfully implemented in music perception studies involving normal-hearing and deaf adults (Karam et al., 2009a,b; Ammirante et al., 2013, 2016; Giordano, 2016; Branje et al., 2014). However, – to our knowledge – no study so far has applied the vibrotactile transduction of music stimuli to infant research. Therefore, the Infant vibrotactile system was developed to adapt the MHC system to the infant body, thus proportionally reducing the number of channels transmitting the signal. Specifically, the Infant vibrotactile system consisted of two channels of vibrotactile transducers embedded in a pillow and placed in the seat-back of an infant high-chair (Figure 4.2).

To transmit stimuli into VTS sensory modality, the original stereo sounds were processed by dividing kick (bass) sounds and snare (high) sounds into two separate channels. Each was processed with a parametric EQ filter to emphasize the vibration of the actuators and to remove resonances. The bass channel was processed with a Low-pass filter (500Hz, Slope: 24dB/Oct, Q=0.75) and a Low-shelf filter (500Hz, Gain: +24dB, Q=1.00). The high channel was processed with the same Low-pass filter, plus a High-pass filter (100hz, Slope: 24dB/Oct, Q=0.75) and a Peak filter (200Hz, Gain: +24dB, Q=0.30). Each channel was then projected to a different portion of the seat-back. The left channel was displayed on the low part of the back and transmitted bass frequencies, whereas the right channel played high frequencies in snare drums to the upper part of the back. This frequency-based signal projection is based on the tonotopic organization of the human cochlea, where frequency-specific hair cell receptors along the basilar membrane are specialized in detecting individual band frequencies (von Békésy and Peake, 1990; Karam et al., 2009b, Figure 4.2). Each channel consisted of a set of two tactile transducers: the upper channel, responsible for the translation of high frequencies, consisted of two voice coils, whereas the bottom channel, responsible for low frequencies, consisted of two bass shakers. All the tactile transducers received the signal from an amplifier connected through a jack cave to a laptop (Figure 4.2). A remote, infrared eye-tracking camera (Tobii X2-60 Eye-Tracker) placed directly below the screen 60 cm away from the participant recorded the eye movements using bright-pupil technology at a sampling frequency of 60 Hz.

Experimental design and procedure

To investigate infant VTS discrimination abilities, this Study implemented a gaze-triggered, familiarization-test paradigm (Figure 4.3). Participants eye movements and pupil dilation were continuously recorded during the experiment, eliciting the trial presentation. At the beginning of each experimental section, a calibration was run. Participants were presented with small, cartoon images appearing on the screen at five locations (i.e., top-left, top-right, center, bottom-left, and bottom-right) together with a piece of cheerful music to capture their attention. Once the calibration was successfully completed, the experiment started. The gaze-triggered procedure implemented in this Study ensured that each experimental trial only began once the infant's gaze was detected as directed toward the AOI on the screen by the eye-tracker and it stopped when infants looked away for more than 2 seconds. Between trials, an attentional getter was presented on the screen to attract the infant's gaze toward the monitor.

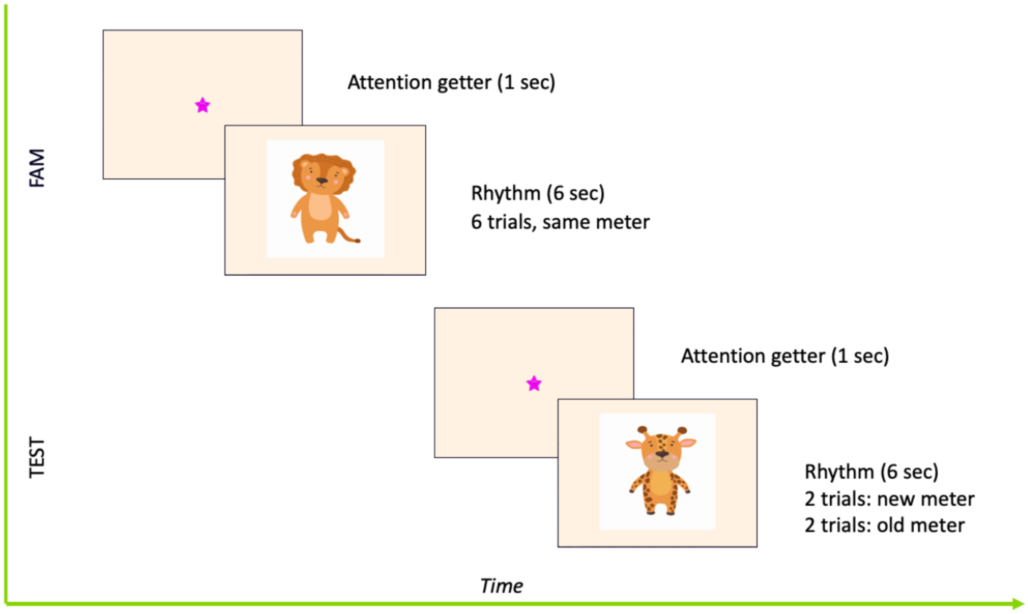


FIGURE 4.3: The gaze triggered, familiarization-test paradigm implemented in Study one. Familiarization phase *FAM* consisted of 3 different VTS rhythms (each lasting for 6 sec) all sharing the same underlying meter and presented twice (6 trials) in randomized order. The test phase (TEST) consisted of a VTS rhythm with a new meter and a VTS rhythm with the familiarized meter both presented twice (4 trials, 6 seconds each) in randomized order.

During the familiarization phase, three different VTS rhythms (6 seconds each) all sharing the same underlying meter were presented twice (6 trials) in random order. Each participant was assigned to a familiarization condition based on the double or triple meter presented during the familiarization phase (Hannon and Johnson, 2005). Immediately after the familiarization phase, the test phase started. Test trials consisted of two different rhythms (6 seconds each) presented twice (4 trials) and in random order, one of which was in double and the other one in triple meter. All infants were presented with the same test trials, regardless of the familiarization conditions. Therefore, test trials with double meter were familiar to the infants from the double familiarization group and novel to infants from the ternary familiarization group; whereas, test trials with triple meter were familiar to infants from the ternary familiarization group and novel to infants from the double familiarization group. During the experimental section, infants were seated on a high chair positioned 60 cm away from the monitor and were tested individually in a quiet room of their kindergarten. Kindergarten's educators familiar to the infants assisted the experimental session to comfort them in case of crying or fussiness. Semi-darkness constant luminance was guaranteed by a lamp positioned 1 m behind the participant.

Statistical analyses

Only data from participants who reached at least 1 valid test trial per type (i.e., novel and familiar) were analyzed. Looking times and pupil dilation displayed toward visual stimuli triggering VTS rhythms were the collected dependent measures as described below.

Looking times data refer to the amount of time that infants spend looking at the images. Since the presentation of VTS rhythms was elicited by the infant's gaze, looking times toward the AOI were collected as an index of attention displayed toward the experimental stimuli (Vouloumanos and Werker, 2004; Russo et al., 2021; Calignano et al., 2021a). In the present Study, looking times were analyzed and interpreted according to familiarization-test studies (Cooper and Aslin, 1990; Jusczyk and Aslin, 1995; Marcus et al., 1999) but also considering the effect of age and individual variability. Specifically, data were analyzed with the open-source R software (R Core Team, 2022) by means of Generalized Mixed-Effects models (Nieuwenhuis et al., 2012; Fox, 2015, GLMM). This statistical approach is well known to adequately deal with methodological issues common to infant research including small sample size, reduced number of observations per trial and participant, and high individual variability due to cognitive and maturational factors (Oakes, 2017). Moreover, GLMM allows for specifying the distribution family, overcoming the ANOVA assumption of normally distributed residuals with uniform variability across the levels of the predictors (Fox, 2015). Accordingly, the gamma family was selected for modelling the looking times data. All models were fitted with the lme4 package (Bates et al., 2014). Starting from a null model with a random intercept for participants, age (in days) as a continuous predictor and then test trial type (i.e., *novel* vs *familiar*) and familiarization group (i.e., *double* or *ternary*) as categorical predictors were gradually inserted in the model.

To find the best approximation to the true model, it was therefore followed a model comparison approach, using the likelihood ratio test (LTR), Akaike Information Criterion (Akaike, 1974, AIC), and AIC weight as indexes of the goodness of fit. Specifically, the former tested the hypothesis of no differences between the likelihoods of two nested models while the AIC and AIC weight together gave information on the relative evidence of models (i.e., likelihood and parsimony). That is, the model with the lowest AIC and the highest AIC weight is preferred (Wagenmakers and Farrell, 2004).

Together with looking times, changes in pupil size under constant luminance were collected and taken as a measure of cognitive processing during stimuli presentation (Beatty et al., 2000; Einhäuser, 2017; Mathôt and Van der Stigchel, 2015; Calignano et al., 2021a,b). Pre-processing steps (i.e., transforming the raw data into a format for statistical analysis and visualization) were performed following the Hepach and Westermann (2016) procedure and the Mathôt and Vilotijević (2022) guidelines (Figure 4.4). Specifically, the library PyGaze (Dalmaijer et al., 2014) was implemented in the software OpenSesame (Mathôt et al., 2012) used to program and administrate the experimental procedure, generating a unique row file (.tsv) per participant logging in all the variables of interest. Row data (.tsv) from each participant were converted into .csv files in Python and then aggregated into a cumulative dataset for the entire sample. Missing data (i.e., data where the eyetracker failed to detect the pupil due to technical or artifacts reasons, indicated with the value 1) were converted into NaN values (not a number). Invalid data (i.e., data where the recorder measure does not reflect the actual pupil size) were removed. Specifically, given that the human pupil size varies between 2 and 8 mm in diameter (Mathôt and Vilotijević, 2022; Pan et al., 2022), values outside this range were excluded from the analysis. Once the coherence between left and right pupils was assessed (R = 0.91), data from the two pupils were averaged in a cumulative variable. Then, baseline correction was performed by subtracting the mean pupil size recorded during a baseline period of 100ms from all subsequent pupil-size measurements per trial and participant.

This technique allows for better detection of the effect of task-evoked responses (Mathôt, 2018, *psychosensory pupil response*) by reducing the effect of luminance (Mathôt, 2018, *pupil light response*) and the impact of trial-to-trial fluctuations while dealing with inter-individual variability. As in the case of looking times, pupil data were analyzed with generalized linear mixed models (GLMM) accounting for both random and fixed effects while specifying the distribution family. Specifically, the gaussian distribution was selected to model the pupillary data (van Rij et al., 2019; Calignano et al., 2021a). All models were fitted with the lme4 package (Bates et al., 2014) in R software (R Core Team, 2022). Similarly to looking times data, a null model with a random intercept for participants was performed first, then age (in days) as a continuous predictor, test trial type (i.e., novel vs familiar), and familiarization group (i.e., double or ternary) as categorical predictors were gradually inserted in the model. To find the best approximation to the true model, we therefore followed a hierarchical step-wise forward model comparison approach (Heinze et al., 2018), based on AIC (Akaike Information Criterion) and delta AIC and AIC weight as indexes of the goodness of fit. Specifically, the model with the lowest AIC and the highest AIC weight was preferred (Wagenmakers and Farrell, 2004).

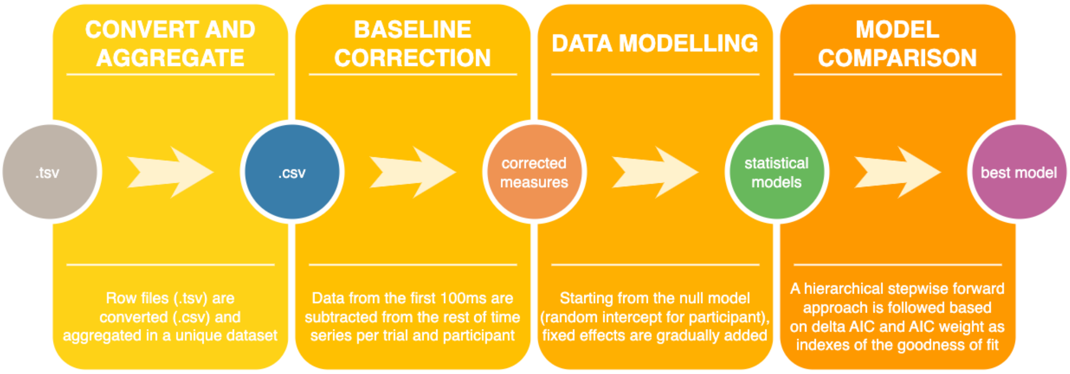


FIGURE 4.4: Workflow of the pupillary data processing. Boxes (from yellow to orange) represent the processing steps while circles (from gray to purple) represent the input and output of each step.

4.1.3 Results

Results will be presented for the looking time data first (Section 4.1.3) and then for the pupillometry analysis (Section 4.1.3).

Looking Times

Looking times from the Familiarization phase (Figure 4.5) indicate a significantly decreasing pattern of looking times per trial (trial 5: b=-1210.79, SE=433.70, t=-2.792, *p*=0.006; trial 6: b=-906.59, SE=429.22, t=-2.112, *p*=0.037). Therefore, this first sanity check confirms that infants succeeded in encoding the stimuli, decreasing their interest with increasing presentations (Cooper and Aslin, 1990; Vouloumanos and Werker, 2004).

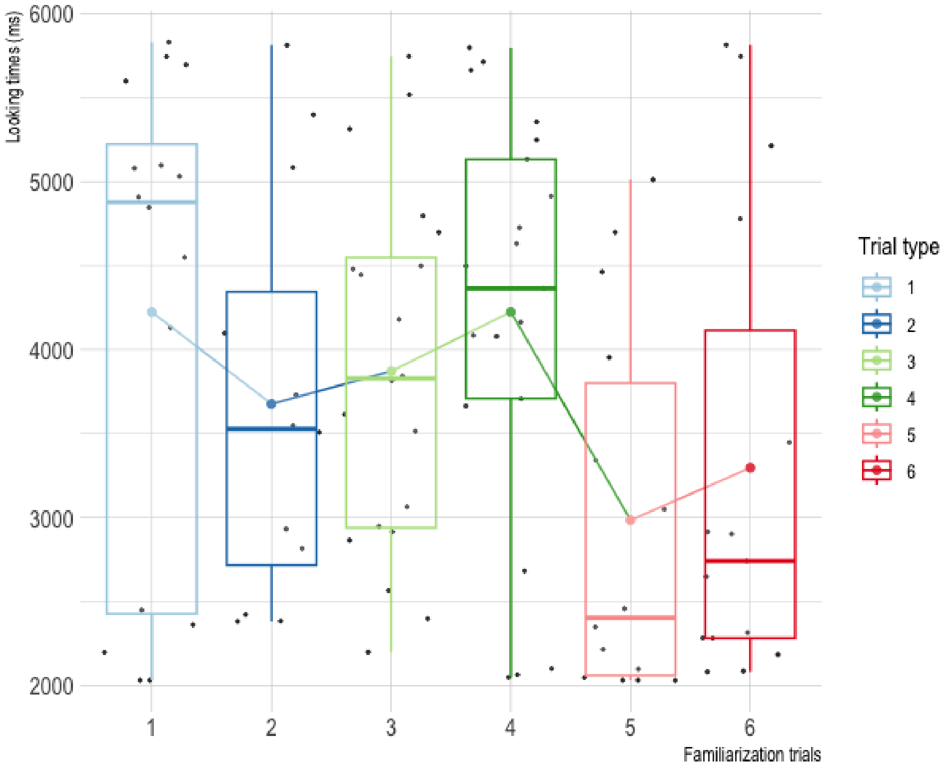


FIGURE 4.5: Looking times (y-axis) for the Familiarization phase (x-axis).

Descriptive statistics for the Test phase (Figure 4.6, left) show the fixation time for trial type (i.e., new/old) across age, for both familiarization groups (i.e., double/triple); together with the number of observations per age (Figure 4.6, right). Specifically, from descriptive statistics it can be seen how infants' looking times change across age with younger infants looking slightly longer toward novel stimuli while older infants seem to display the opposite pattern (Figure 4.6, left). Overall, older infants seem to display an increased difference in looking times directed toward novel vs familiar trials compared to young infants. Moreover, both younger and older infants seem to show the same pattern of results despite the familiarization group, although younger infants familiarized with double meters tend to look toward both novel and familiar trials longer.

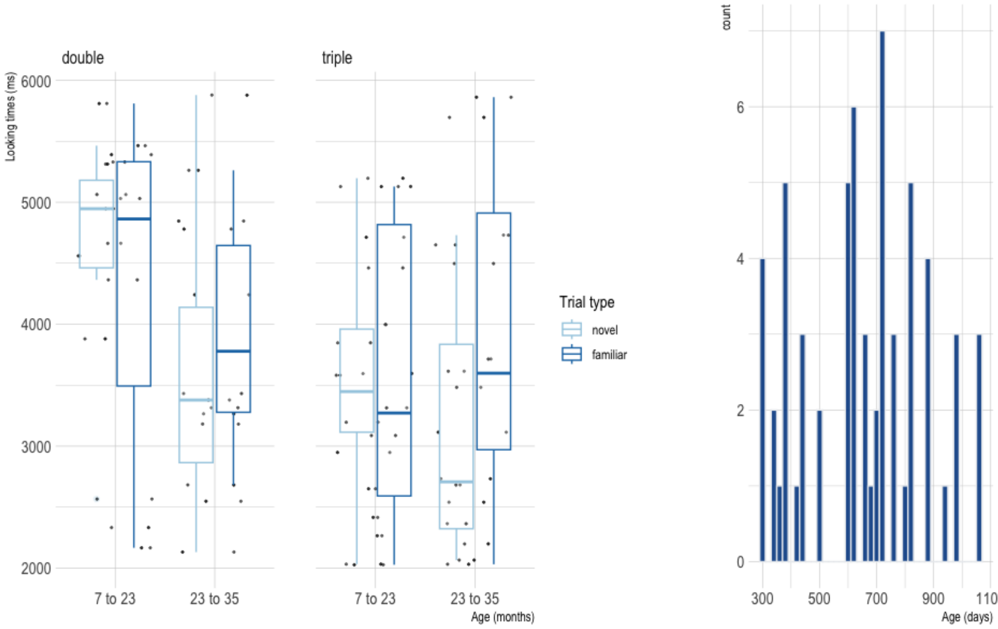


FIGURE 4.6: Descriptive statistics of looking time data. Fixation times for trial type (novel/familiar) across age groups (7 to 23 / 23 to 35 months) in each familiarization group (double/triple; left). Number of observations (i.e., sampled values per trial and participant) per each age (in days; right).

As shown in Table 4.1, data were best explained by Model 2 (Wagenmakers and Farrell, 2004, lower dAIC and higher AICw). Specifically, Model 2 accounted for the additive effects of trial type and familiarization group on looking times (Figure 4.7). However, only the familiarization group effect was significant, with longer fixation times performed by infants familiarized with a double meter (b=-0.20, SE=0.09, t=-2.14, *p*=0.0317).

TABLE 4.1: GLMM comparison for looking times in Study 1

Models	Deviance	dAIC	AICw
M.0 Looking times ∼(1 ∣ id)	3.7664	0.22	0.29
M.1 Looking times ∼trial type + (1 ∣ id)	3.6885	1.34	0.16
M.2 Looking times ∼trial type + familiarization group + (1 ∣ id)	3.6215	0.00	0.32
M.3 Looking times ∼trial type + familiarization group + age + (1 ∣ id)	3.6206	2.44	0.09
M.4 Looking times ∼trial type * familiarization group + age + (1 ∣ id)	3.6299	5.07	0.03
M.5 Looking times ∼trial type * age + familiarization group + (1 ∣ id)	3.4080	2.25	0.10
M.6 Looking times ∼trial type * age * familiarization group + (1 ∣ id)	3.3911	10.48	0.00

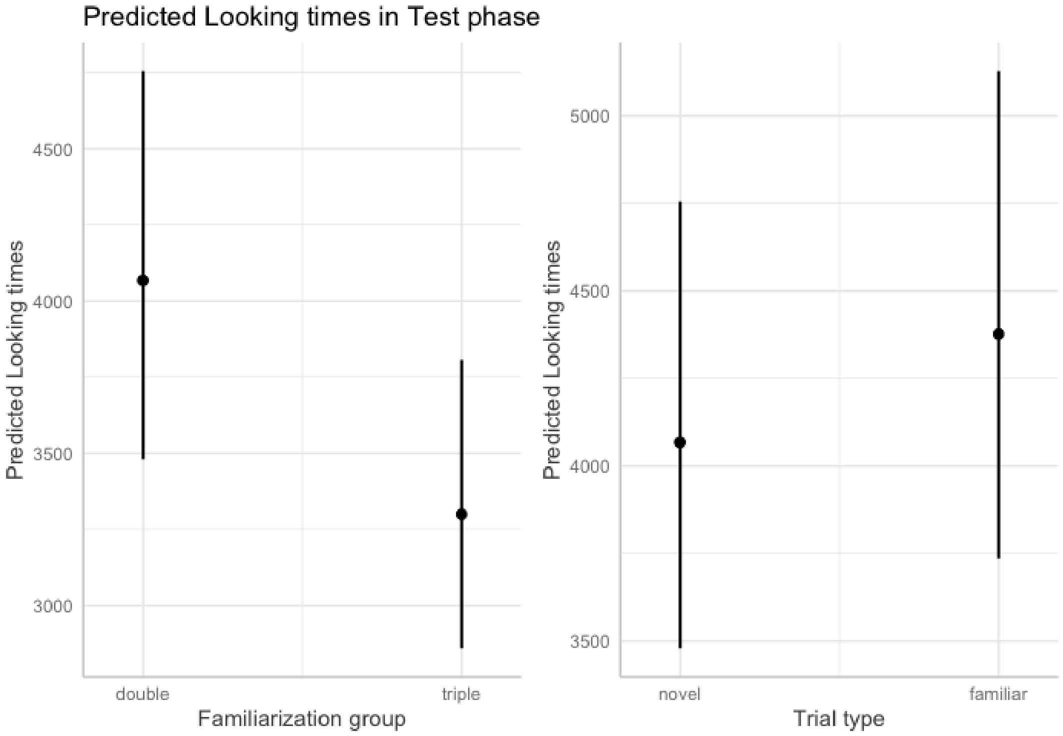


FIGURE 4.7: Effect plots of the best model (M2) predicting looking times in the test phase. Predicted looking times for novel vs familiar trial type (left) and for double vs triple familiarization group (right).

Pupillometry

The following descriptive statistics show the pupillary data for trial type (i.e., new/old) across age (Figure 4.8, left) in both familiarization groups (i.e., double/triple) together with the number of observations per age (Figure 4.8, right). From this plot, it can be seen how data change across age. Specifically, younger infants seem to show a slight increase in pupil diameter in response to familiar stimuli while older infants seem to display an opposite trend (Figure 4.8, left). Moreover, older infants familiarized with a double meter seem to show higher increases in pupil diameter compared to other infants. Lastly, it must be noted that the number of observations is substantially higher with pupillary data compared to looking times data (Figure 4.8, right). In fact, changes in pupil diameter are coded as a continuous time series of pupil diameter (in mm) sampled approximately each 17ms by a 60Hz eye-tracker. This results in a much higher number of observations per trial and participant (3000) compared to the discrete values of looking time data (i.e., a single value per trial and participant).

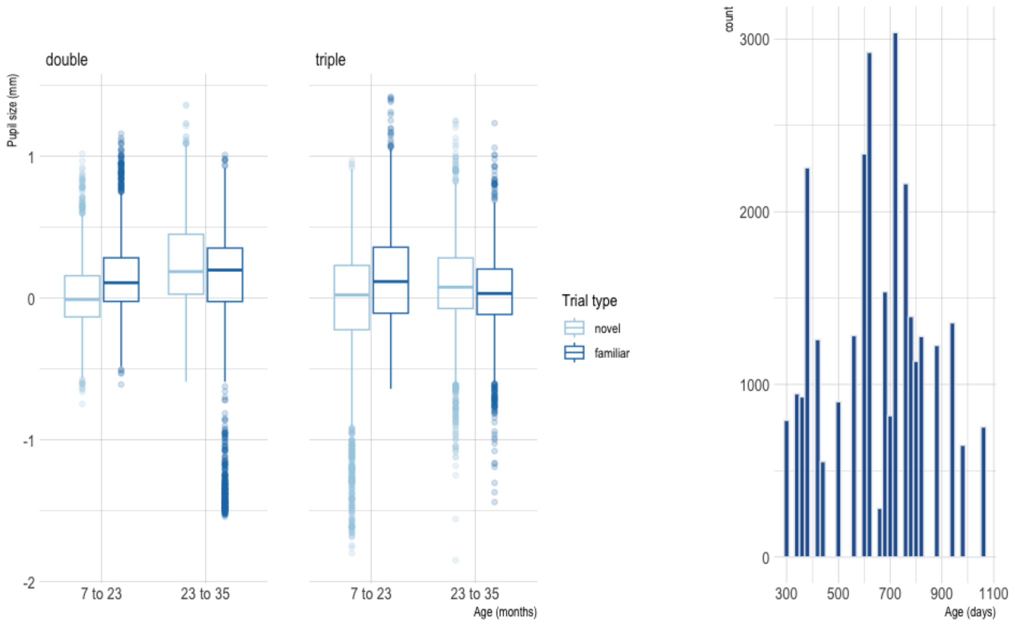


FIGURE 4.8: Descriptive statistics of pupillary data. Baseline-corrected pupil diameter for trial type (novel / familiar) across age (7 to 23 / 23 to 35 months) in each familiarization group (double/triple; left). Number of observations (i.e., pupillary diameter in mm sampled at Hz) per age (in days; right).

In modelling pupillary data, the higher number of observations of this variable allowed for including age as a continuous variable in the models thus closer exploring the changes in pupil diameter across infancy. As shown in Table 4.2, data were best explained by Model 5 (Wagenmakers and Farrell, 2004, lower dAIC and higher AICw). Specifically, Model 5 accounted for multiple interactions across age, trial type, and familiarization group effects on pupillary data (Figure 4.9).

TABLE 4.2: GLMM comparison for pupillary data in Study 1

Models	Deviance	dAIC	AICw
M.0 Pupil diameter ∼(1 | id)	16580	346.69	0
M.1 Pupil diameter ∼age + (1 | id)	16595	364.29	0
M.2 Pupil diameter ∼age + trial type + (1 | id)	16595	366.03	0
M.3 Pupil diameter ∼age + trial type + familiarization group + (1 | id)	16323	95.67	0
M.4 Pupil diameter ∼age * trial type + familiarization group + (1 | id)	16323	98.38	0
M.5 Pupil diameter ∼age * trial type * familiarization group + (1 | id)	16219	0.00	1

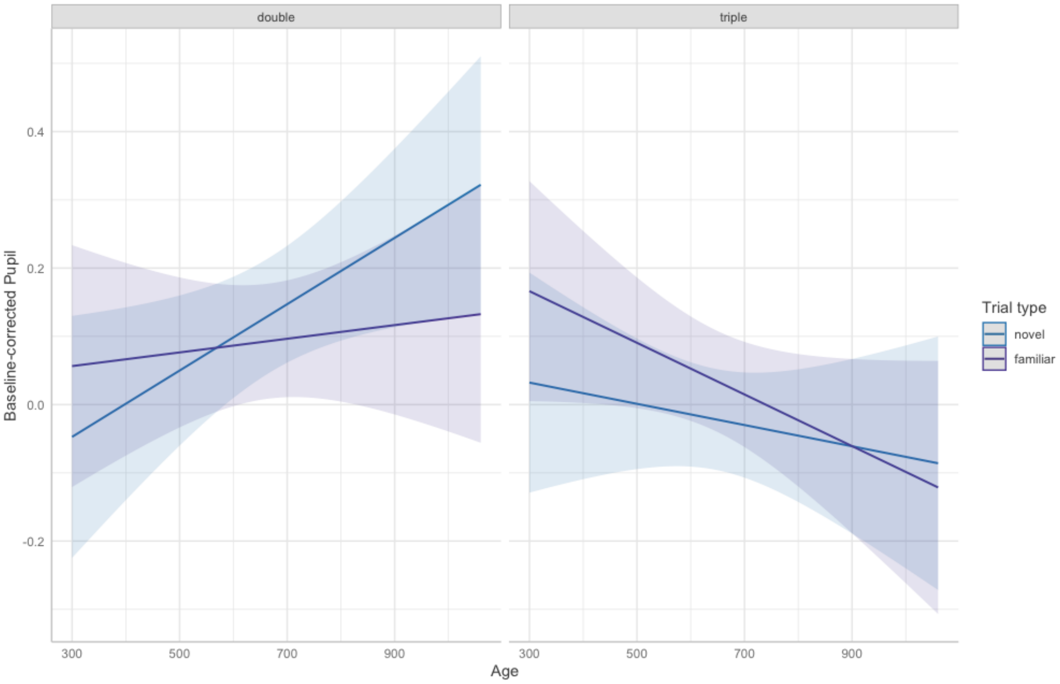


FIGURE 4.9: Effect plots of the best model (M5) predicting pupil size in the test phase. Predicted changes in pupil dimeter for novel vs familiar trial type (left) and for double vs triple familiarization group (right).

Specifically, the single effect of trial type (b=0.219, SE=0.02, t=10.650, p<0.001) and age (b=0.0004, SE= 0.0002, t=2.233, p=0.026) were significant as well as their interaction (b=-0.0003, SE= 0.00002, t=-13.198, p<0.001). Moreover, the familiarization group and age also show a significant interaction (b=0.0006, SE=0.0003, t=-2.118, p=0.034). Lastly, the three-way interaction of all predictors resulted as significant (b=0.0001, SE=0.00004, t=4.007, p<0.001) as shown in Figure 4.9.

4.1.4 Interim discussion

This Study investigated rhythm discrimination under vestibular-tactile-somatosensory (VTS) modality in infancy. To this end, an Infant vibrotactile device for music transduction was designed based on previous works on sensory substitution system in music technology (Karam et al., 2009a,b; Giordano, 2016; Branje et al., 2014). This system was then combined with an eye-tracker to implement a gaze triggered, familiarization-test paradigm. Based on the performed statistics, the best model (M2; see Paragraph 4.1.3) predicting looking times in this Study accounted for the additive effect of trial type and familiarization group, net of inter-individual variability. Specifically, looking times toward contingent visual stimuli seem to increase for VTS rhythms with familiar meters, but this increase does not reach the statistical significance (see Figure 4.1, right). Only the familiarization group effect reached significance. This means that infants familiarized with double VTS rhythms looked significantly longer in the test phase compared to infants familiarized with a triple meter. Since the test phase was the same for all participants (see Paragraph 4.1.2), this difference can be attributed to the effect of the familiarization phase. That is, familiarizing infants with double meters produced a significant increase in looking times during the test phase compared to infants familiarized with triple meters. This result is in line with literature showing that infants better detect changes and reproduce melodic as well as irregular harmonic sequences with double meters (Bergeson and Trehub, 2006; Bhatia et al., 2013; Drake and Botte, 1993). This phenomenon has been explained referring to the prenatal experience with rhythmic stimuli endogenous to the intrauterine environment.

Specifically, it has been proposed that rhythms displaying a double patterning (e.g., maternal heartbeat, respiratory cycles) might provide the foetus with its first sense of pulse already in the womb (Ullal-Gupta et al., 2013; Teie, 2016). As consequence, double rhythmic patterns perceived in the womb might have a first impact on the development of neural and physiological systems supporting rhythm development, thus shaping early discrimination abilities toward familiar patterns (DeCasper and Sigafoos, 1983; Doheny et al., 2012; Lahav et al., 2007; Tichko et al., 2022).

Furthermore, all the infants tested in this Study were Italian-native speakers. Accordingly, double meters are more prevalent in Western countries than compound and ternary meters (Huron, 2006). Effects of musical enculturation were found in cross-cultural studies showing difficulties in perceiving unfamiliar metrical patterns in Balkan and Turkish music in Western infants (Hannon and Johnson, 2005; Hannon and Trehub, 2005; Kalender et al., 2013; Soley and Hannon, 2010; Hannon et al., 2012). Paralleling the findings from language and faces processing (Maurer and Werker, 2014), music enculturation was proposed as a perceptual narrowing process involving the progressive attunement to culturally relevant categories of stimuli such as double meters for Western listeners (Lewkowicz, 2014; Maurer and Werker, 2014). Accordingly, musical enculturation processes were found to begin in infancy and continue until adulthood (Yates et al., 2017; Trainor and Hannon, 2013). In the present Study, looking data showed a trend consistent with this literature (see Figure 4.6). Specifically, older infants (i.e., 23 to 35 months) seem to display longer fixation times for familiar meters showing finer discrimination abilities compared to younger participants (7 to 23 months). Moreover, previous studies with a similar design (Hannon and Johnson, 2005) reported increased looking times for novel meters in 7-month-old infants, which is in line with the present data displaying a comparable trend (see Figure 4.7, left).

However, according to the performed statistics, the trial type effect did not reach significance. Therefore, the aforementioned considerations must be taken with caution and further research is needed to confirm the observed trend. Regarding the lack of significance, it must be noticed that the number of observations per trial and participant across ages was considerably restricted (see Figure 4.6). In fact, looking times are a behavioral measure consisting of the accumulated amount of time (in ms) for fixations directed toward a visual stimulus. Therefore, each trial consists of a single, discrete observation. Moreover, the inclusion criteria for statistical analyses (see Paragraph 4.1.2) were to reach a minimum of 1 observation per trial type in the test phase. Given a minimum of 2 and a maximum of 4 observations per participant, the final sample size resulted in 62 total observations. Accordingly, it was necessary to lower the number of points per axis to evaluate the adaptive Gauss-Hermite approximation to the log-likelihood, that is a faster but less accurate setting (nAGQ=0) to estimate convergence (Bates et al., 2014; Lindstrom and Bates, 1990).

Therefore, the nature of this measure calls for a bigger sample size to increase the number of observations and thus perform strong statistical approaches such as generalized linear mixed models or – even better – Bayesian statistics. Although this number of observations was sufficient to estimate the parameters of M2 accounting for the additive effect of trial type and familiarization group, the number of parameters in the following models might have been too high to be estimated from 62 observations. Therefore, despite the limitations related to the nature of a behavioural measure, an interesting trend emerged from the data as visible and similar across groups. Thus, the possibility for this trend to reach significance with an increased number of observations remains open. In summary, a larger sample of observations is needed to effectively estimate fixed (i.e., trial type and familiarization group) and random (i.e., individual variability) effects across age in infant experimental designs involving looking times data.

By contrast, pupillometry data do offer a significantly larger number of observations per trial and participant. In fact, the sampling rate of commonly used eye-trackers ranges from 60 to 120 Hz. In the present Study, a Tobii X2-60 eye-tracker with a sampling frequency of 60 Hz allowed for collecting data on the pupil diameter every 17 ms. With the same sample size, this led to time series of about 3000 observations per trial and participant (Figure 4.8). Therefore, pupillary data allowed for designing models estimating both random and fixed effects without penalizing the log-likelihood adaptation (Bates et al., 2014; Lindstrom and Bates, 1990, nAGQ=1). This resulted in the estimation of the best model (M6) accounting for a complex interaction between age as a continuous predictor and trial type and familiarization group as categorical ones, net of individual variability. Specifically, this model predicted an increase in pupil diameter for familiar meters in younger infants while the opposite trend emerged from the older ones (Figure 4.9). This pattern was approximately the same in both familiarization groups. Moreover, while pupillary increases predicted by M6 are comparable within younger infants in the two familiarization groups, it might be noticed that differences for novel and familiar stimuli increase with age for infants familiarized with double meters and decrease for infants familiarized with triple ones.

Given that differences displayed toward novel vs familiar test stimuli are taken as a measure of discrimination abilities in infant studies (Cooper and Aslin, 1990; Vouloumanos and Werker, 2004), increased discrimination abilities might be inferred from the present data for infants familiarised with double meters. Paralleling looking time data, this result is in line with previous findings on musical enculturation processes in development (Trainor and Hannon, 2013). Accordingly, following the principles of perceptual narrowing, processing abilities in young infants are expected to be broader and to tune toward culture-specific signals along development; at the same time, processing abilities for stimuli uncommon to native environments are expected to decrease (Lewkowicz, 2014; Maurer and Werker, 2014). This appears to be the trend displayed in this Study, confirmed by the accordance of behavioral (i.e., increased looking times for infants familiarized with double meters) and physiological (i.e., increased pupil dilation for familiar over novel stimuli in infants familiarized with double meters) measures. Therefore, results from this Study supports age-related changes in rhythmic abilities in infancy.

4.1.5 Conclusions to Study one

This Study aimed at testing whether infant rhythmic abilities encompass stimuli perceived via vestibular-tactile-somatosensory (VTS) modality. Based on the results, this Study might provide a first positive answer to this question. In fact, even though further research is needed to replicate and extend these results, the updated techniques from human-computer interaction (Karam et al., 2009b, sensory substitution systems and computational sonology), infant research (Hepach and Westermann, 2016; Mathôt and Vilotijević, 2022, eye-tracking and gaze-triggered designs), and statistics (Wagenmakers and Farrell, 2004, GLMM and model comparison techniques) applied in this Study brought new, promising evidence on rhythm perception across sensory modalities in infancy. Specifically, this evidence might help in clarifying the extent of sensory abilities for rhythmic stimuli thus contributing to both experimental and applied research. In fact, exploring how infants process rhythm can shed light on the perceptual abilities underlying complex cognitive processes as well as the extent to which these processes might be considered as general abilities serving multiple cognitive functions. Accordingly, basic mechanisms underlying rhythm processing (Fiveash et al., 2021, the PRISM framework, see Chapter 3) are hypothesized to be crucial to language development.

Furthermore, rhythmic difficulties were found across neurodevelopmental disorders (Lense et al., 2021) pointing out atypical rhythm as a possible risk factor for neurodevelopmental disorders (Ladányi et al., 2020, the ARRH, see Chapter 3). Finally, given the crucial role of rhythm in cognitive development, exploring rhythmic abilities across different sensory modalities can be further informative for designing screening and intervention programs for early sensory deprived infants as well as children with sensory-related issues (Tomchek et al., 2014; Ghanizadeh, 2011, e.g., ASD and ADHD). For the aforementioned reasons, the next step of this work was to investigate whether the VTS rhythmic abilities measured in this first Study were related to early language development. Specifically, the same cohort of infants participating in Study one was tested with a linguistic task in Study two; then, an exploratory analysis investigated the relationship between individual performances across the two experiments. The linguistic task and its link to VTS rhythmic abilities are described in the next section.

4.2 Study two: linking VTS rhythms and language

In this Study, the emerging linguistic abilities of infants from 7 to 35 months will be tested through a discrimination task for novel object-label pairs. Specifically, prosodic and phonological abilities will be measured by investigating how infants react to mispronunciations, occurring at the stress or the phonological level, in recently familiarized labels. Discrimination abilities will be inferred by physiological (i.e., pupillometry) measures collected through the eye-tracker for screen images. Specifically, infants will be presented with a visual stimulus (i.e., a cartoon character) and a corresponding auditory label (i.e., a disyllabic unit like /bosa/). Labels are presented in their original form (i.e., familiar; F condition) or altered in their prosodic (i.e., novel stress; NS condition) or phonological (i.e., novel phonological; NP condition) structure. Changes occurring in pupil diameter are expected across conditions and taken as measures of prosodic and phonological discrimination abilities. Crucially, the same cohort of infants participating in Study one was tested again in Study two, since the main aim of this investigation was to explore the link between rhythmic and linguistic abilities in infancy. This aim has been addressed by means of a within-participants design. As for Study one, data collection was made possible thanks to the contribution of 5 kindergartens as part of the Cooperative Now Project. Results were discussed with researchers from University of Padua and abroad during meetings, research stays, conferences, and project presentations.

4.2.1 Aim and research questions

The aim of this Study was to test the emerging linguistic abilities (i.e., phonological and prosodic processing) underlying lexical encoding of label-object pairs in infancy. Then, the link between linguistic abilities (Study one) and the previously measured VTS rhythmic skills (Study two) was investigated by means of an exploratory analysis with the aim to test whether one domain might be informative about the other. The motivation for this Study comes from the literature on early language showing that lexical encoding is sensitive to feature alterations (Durrant, 2017; Swingley and Aslin, 2000; Yoshida et al., 2009; Delle Luche et al., 2015). Specifically, early word representations seem to be characterized by a greater phonological specificity than expected (Swingley and Aslin, 2000; Mani et al., 2012).

Consistently, infants show fine phonological abilities since the first stages of language acquisition (Curtin and Archer, 2015; Velleman and Vihman, 2007) and later in development, as their experience with the native language grows (Kuhl et al., 2006). Similarly, basic prosodic skills are already in place at birth (Gervain, 2018; Abboub et al., 2016) while fine discrimination abilities emerge by experience with native prosodic patterns later in development (Jusczyk et al., 1999; Johnson and Jusczyk, 2001, see Chapter 3).

Crucially, most of the studies investigating word representation in infancy rely on target looks in looking while listening tasks (Hohle et al., 2006; Mani et al., 2012; Swingley and Aslin, 2000; White and Morgan, 2008; Delle Luche et al., 2015). In this paradigm, an auditory label is presented while two images (i.e., a target and a distractor) appear on the screen. Looks toward the target are expected to decrease in case of mispronunciations (Mani and Plunkett, 2011; White et al., 2013). Even though this paradigm offered useful insights into early word acquisition, it also presents some limitations. For instance, the use of a behavioural measure (i.e., looking times) can be sensitive to variables other than the one of interest (DePaolis et al., 2016; Roder et al., 2000; Soley and Hannon, 2010; Oakes, 2017, e.g., motivation, familiarity, maturational level). Furthermore, the presence of two simultaneous objects might affect the results due to familiarity and / or spontaneous preference (White and Morgan, 2008; Tamási et al., 2019). Therefore, the present Study investigated object-label encoding by analysing variations in pupil diameter while presenting an auditory label and a unique visual referent. As stated in Study one, pupillometry is not only advantageous since represents a continuous measure, but it's also a physiological measure less sensitive to conscious factors such as motivation, preference, or maturation (Klingner, 2010). As a direct measure of cognitive load rather than preference, changes in pupil diameter can indeed accurately reflect the effort of encoding mispronounced labels than looking time or preferential looking tasks. Therefore, pupillometry might provide a detailed index of early object-pairs representation in infancy. Previous studies demonstrated that pupil diameter is sensitive to mispronunciations, with an increase in deviant labels compared to corrected or unrelated ones (Fritzsche and Höhle, 2015).

Interestingly, the continuous and time-locked nature of the pupil response allowed for studying whether infants are sensitive to different degrees of mispronunciations (Tamási et al., 2019, 2017). Specifically, infants at 35 months of age have been found to display a larger dilation in response to labels that deviate more than ones that deviate less (Tamási et al., 2019, 2017). Therefore, pupillometry is an informative window into early lexical abilities in infancy.

In the present Study, variations in pupil diameter to a visual referent while listening to corrected or altered (for phonological or prosodic features) labels are collected to replicate previous findings on pupil sensitivity to different degrees of mispronunciations in infancy. Moreover, infants from 7 to 35 months will be tested to investigate the developmental trajectory leading to the fine word representations reported in previous studies (at 35 months). Lastly, relationships between linguistic skills and VTS rhythmic abilities (from Study one) will be explored to test whether early rhythmic and linguistic abilities are related in infancy. According to the literature showing a core link between rhythm and language in infancy, a positive relationship between the two studies was expected. Implications regarding the extent to which the emergence of linguistic skills in infancy is related to early rhythmic abilities across modalities will be discussed, with important insights for language screening and training practices throughout development.

4.2.2 Method

Participants

The same cohort of forty-five infants (21 females, M age = 661.6 days, SD = 192.6) participating in Study one also participated in this Study, recruited from kindergartens from Padua (IT) and adjacent municipalities as part of the Cooperative Now Project (www.coopnow.it. The inclusion criteria for all participants were: to be in good health, to have no sensory or neurological disorders or any familiar language disorder, and to be native Italian speakers. Five participants were excluded from the analysis because of fussiness/crying (n=4) or because data did not meet the inclusion criteria set prior to data collection of at least 1 test trial per condition (n=1). The final sample included 40 participants. Lastly, a sample of thirty-two participants successfully completed Study one and Study two and was therefore included in the exploratory analysis between the two tasks. The research protocol was approved by the Ethics Committee of the Human Inspired Technology research centre, University of Padua (protocol number: 202066R1). The research was conducted in accordance with the principles of the Declaration of Helsinki. Caregivers signed a consent form before taking part in the experiment while verbal consent was asked to participants before starting the experimental section.

Stimuli

Three different labels were selected from the Novel Object and Unusual Name database (Horst and Hout, 2016, NOUN). This online and free database offers a collection of unusual names (i.e., pseudo-words and novel nouns) for psychological research. The three items selected for this Study were three disyllabic pseudo-words with a CVCV sequence (see Table 4.3). The three pseudo-words were presented in their original form (familiar; F condition) or altered for phonological (novel phonological; NP condition) or prosodic (novel stress; NS condition) features. This resulted in the 9 items reported in Table 1. Each label was paired with a unique visual reference. Images consisted of three cartoon characters and were presented on a neutral background on the monitor. The image areas corresponded precisely to the areas of interest of the eye-tracker (AOI), which measured 10 × 10 cm (9.554 deg) and remained visible throughout the entire trial. Visual stimuli were selected from the online free database From: (freepik.com and were equated for to reduce luminance-induced variability in the pupillometry measurements (Mathôt and Vilotijević, 2022; Hepach and Westermann, 2016, see Figure 4.4).

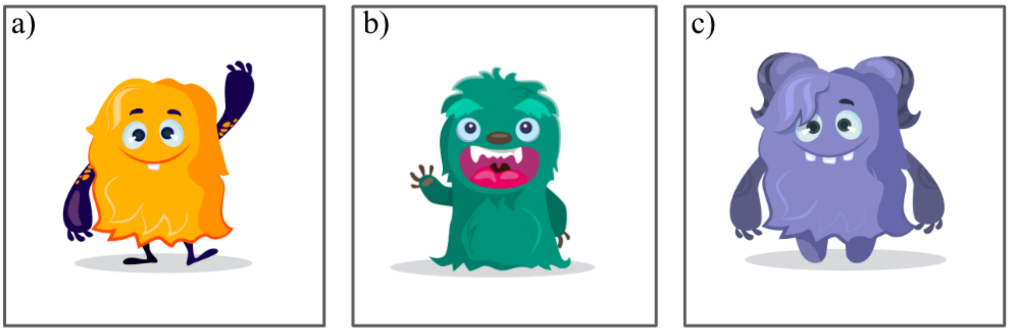


FIGURE 4.10: Visual stimuli for Study two. Each of the three images was a unique referent for the object-label pairs presented in this study. Specifically, the first cartoon character was paired with the label *Loche*, the second one with the label *Nare*, and the third one with the label *Bosa*.

TABLE 4.3: Linguistic stimuli for Study 2

Label	Familiar (F)	Novel Phonological (NP)	Novel Stress (NS)
Bosa	*Bosa*	*Mosa*	*Bosà*
Loche	*Loche*	*Lome*	*Loché*
Nare	*Nare*	*Pare*	*Naré*

Apparatus

The experiment was programmed and presented through the Open Sesame software 3.1 (Mathôt et al., 2012) running on a computer laptop (Acer travel mate 5772 g). Visual stimuli were displayed on a 27-inch monitor (Philips 300 × 300). Labels were auditory presented through two loudspeakers placed on both sides of the monitor. A remote, infrared eye-tracking camera (Tobii X2-60 Eye-Tracker) placed directly below the screen, 60 cm away from the participant, recorded the eye movements using bright-pupil technology at a sampling frequency of 60 Hz.

Experimental design and procedure

To investigate discrimination abilities, this Study implemented a gaze-triggered, familiarization-test paradigm (Figure 4.11). Variations in pupil dilation were continuously recorded during the experiment, eliciting the trial presentation. At the beginning of each experimental section, a calibration was run. Participants were presented with small, cartoon images appearing on the screen at five locations (i.e., top-left, top-right, centre, bottom-left, and bottom-right) together with a piece of cheerful music to capture attention. Once the calibration was successfully completed, the experiment started. The gaze-triggered procedure implemented in this Study ensured that each experimental trial only began once the infant's gaze was detected within the AOI by the eye-tracker and it stopped when infants looked away for more than 2 seconds. Between trials, an attentional getter was presented on the screen to attract the infant's gaze toward the monitor.

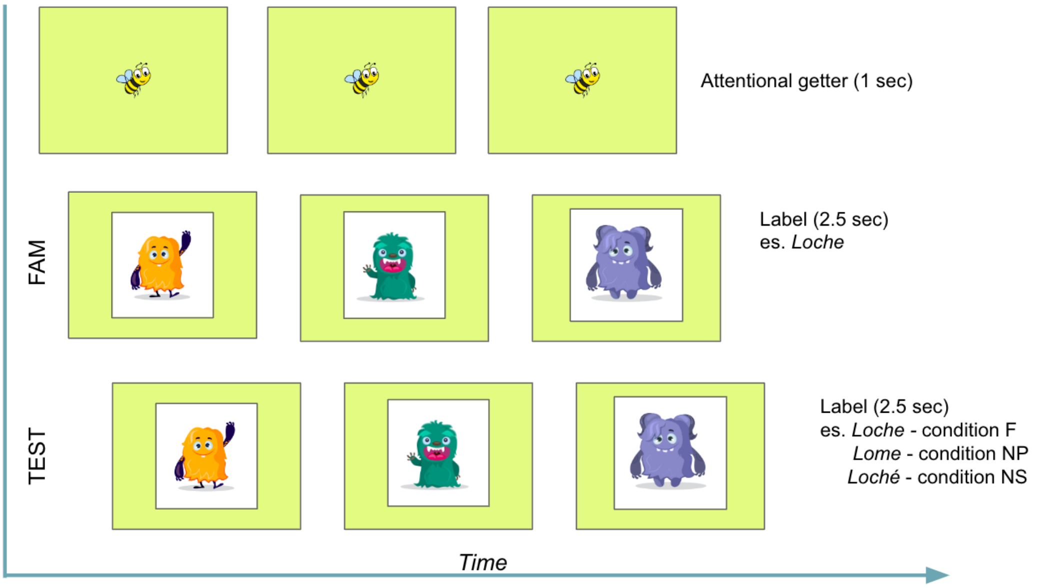


FIGURE 4.11: The gaze triggered familiarization-test paradigm implemented in Study two. Each trial consisted of a familiarization (FAM) and a test (TEST) phase. Familiarization consisted of a label repeated two times (2.5 sec). The test phase (TEST) immediately followed and consisted of a label presented two times (2.5 sec): the first time in its original form and the second time in the familiar, novel phonological or novel stress condition. Between trials, an attentional getter (1 sec) was presented to attract attention toward the monitor.

Each trial consisted of familiarization and a test phase. Before the label presentation, a silent time window of 500ms allowed the eye-tracker to collect baseline data on pupil diameter prior to stimuli presentation. Once the gaze of participants was detected as inside the AOI and registered continuously for 500 ms, the familiarization phase began. In the familiarization phase, a label was repeated two times. With an inter-stimulus interval of 1 sec, the test phase followed and consisted of a label presented two times: the first time in its original form and the second time in the familiar, novel phonological or novel stress condition. Each of the three labels was paired with one of the three visual referents (see Figure 4.3) which remained visible for the entire trial duration (5 sec). Therefore, each trial consisted of one label presented through the familiarization and test phase. Each label-object pair was presented twice in the familiar condition, once in the novel phonological, and once in the novel stress conditions. Therefore, the experimental section consisted of 12 trials: 6 familiar and 6 mispronounced, all presented in random order. This design applied also in previous studies (Russo and Valenza, 2021; Calignano et al., 2021a) allows for testing multiple labels while keeping intact the structure of familiarization-test paradigms. Indeed, every object-label pair is presented 3 consecutive times in its original form in each trial before the test phase presentation. Moreover, this design allows for appropriate pupil data collection since every trial starts with its baseline and every test event is time-locked within a specific time window, facilitating pupillometry analysis. During the experimental section, infants were seated on a highchair positioned 60 cm away from the monitor and were tested individually in a quiet room of their kindergarten. Kindergarten educators familiar with the infants assisted in the experimental session to comfort them in case of crying or fussiness. Semi-darkness constant luminance was guaranteed by a lamp positioned 1 m behind the participant.

Statistical analyses

Only data from participants who reached at least 1 valid test trial per condition (i.e., fam, NS, NP) were analyzed. Variations in pupil diameter displayed toward object-label pairs were collected as the dependent variable. Specifically, changes in pupil size under constant luminance were taken as a measure of cognitive processing during stimuli presentation (Beatty et al., 2000; Einhäuser, 2017; Mathôt and Van der Stigchel, 2015; Calignano et al., 2021b).

Pre-processing steps (i.e., transforming the raw data into a format for statistical analysis and visualization) were performed following the Hepach and Westermann (2016) procedure and the Mathôt and Vilotijević (2022) guidelines (Figure 4.4). Since the pre-processing procedure here applied is the same as Study one, please refer to Section 4.1.2 for further specifications.

Statistical analyses were performed by means of a time-course analysis of pupil dilation. Specifically, to explore the changes occurring in pupil dilation across time, baseline-corrected pupil data were modeled with Generalised Additive Mixed-effect models (GAMM) by using the mgcv package (Wood and Part, 2007, v1.8-33) in R software (R Core Team, 2022). GAM models address the likely non-linear relationship between pupil size variation and time. Data were modeled by using a maximal random structure: the trial was considered as a minimum statistical unit and set a minimum of 20 knots as the maximum number of turning points to be used during the smoothing process (Baayen et al., 2017). To explore whether the experimental manipulations statistically influenced pupil size across time windows, the model's estimates of the differences between conditions were evaluated and visually inspected by means of the itsadug and MuMIn packages (van Rij et al., 2019). Specifically, the itsadug package was used to plot the difference curve based on the model predictions, thus specifying the time windows where differences were expected to be significant as well as the estimated effect size. Since GAM modeling is a nonlinear regression analysis particularly useful for time-course data, the first application of this statistical analysis was performed by setting conditions (i.e., F, NP, NS) as a categorical predictor of pupil variations across time. Moreover, random effects were included as random smooths for participants, trials, and age (Model 1). However, given that variations across the age range were one of the research questions motivating this study, a second model was performed within the test time-window by setting conditions (i.e., F, NP, NS) as a categorical predictor of variations in pupil size occurring across the age continuum (in days). In this second model, random effects were therefore included for participants and trials only (Model 2).

The last step was to test whether performances in the linguistic test were related to the rhythmic abilities measured in Study one. To this aim, a performance index per participant was computed by calculating the proportion of the averaged increment in pupil diameter for novel stimuli in the test phase [novel / (novel + familiar)]. Then, changes in pupil diameter occurring in the linguistic task were modeled by means of Generalized Mixed-Effects models (Nieuwenhuis et al., 2012; Fox, 2015) adding performance indexes in the rhythmic task as a continuous predictor. Model comparison was then performed using the likelihood ratio test (LTR), Akaike Information Criterion (Akaike, 1974, AIC), and AIC weight as indexes of the goodness of fit. To investigate how the relationship between linguistic and rhythmic abilities change across age, an exploratory Generalized Additive Mixed-effect model (GAMM) was performed. Specifically, performance indexes across tasks (i.e., rhythmic, phonological, and prosodic scores) were modeled by adding age as a continuous predictor.

4.2.3 Results

Results for the linguistic task are here reported followed by the age effect analysis in paragraph 4.2.3. Then, the relationship between linguistic and rhythmic performance as well as the variations occurring across age are explored in paragraph 4.2.3.

Linguistic Task

The following descriptive statistics show the pupillary data for conditions in the test phase (i.e., familiar, novel phonological, and novel stress) across age (Figure 4.12, left) together with the number of observations per age (Figure 4.12, right up). Specifically, the test phase consisted of the time window (from 4s to 5s) where labels were pronounced based on conditions (Figure 4.12, right down). From this plot, a slight increase in pupil diameter was registered for labels altered for phonological features in younger infants and this pattern seems to increase with age. By contrast, familiar and labels altered for prosodic features seem to be more stable across age.

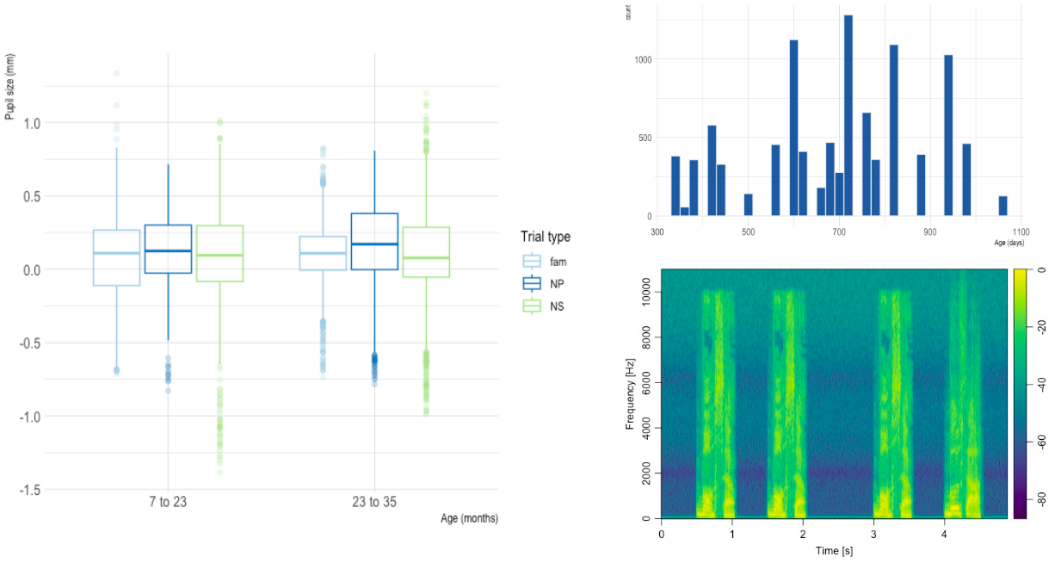


FIGURE 4.12: Descriptive statistics of pupillary data for Study two. Baseline-corrected pupil diameter for condition (i.e., Trial type): familiar (fam), novel phonological (NP), and novel stress (NS) across age (7 to 23 / 23 to 35 months; left). Number of observations (i.e., pupillary diameter in mm sampled at Hz) per age (in days; right up). Spectrogram (Frequency in Hz over Time in seconds) of a trial (right bottom).

For statistical analyses of the linguistic task, pupillary data were modeled by means of GAM models (Wood and Part, 2007). Specifically, the effect of condition on changes in pupil diameter was analyzed in the time dimension, accounting for participants, trials, and age as random effects. The model structure resulted as follows:

M1: *Pupil diameter   trialtype + s(time, by trialtype, k=20) + s(time, age, bs='fs', m=1) + s(time, id, bs='fs', m=1) + s(time, trial, bs='fs', m=1)*

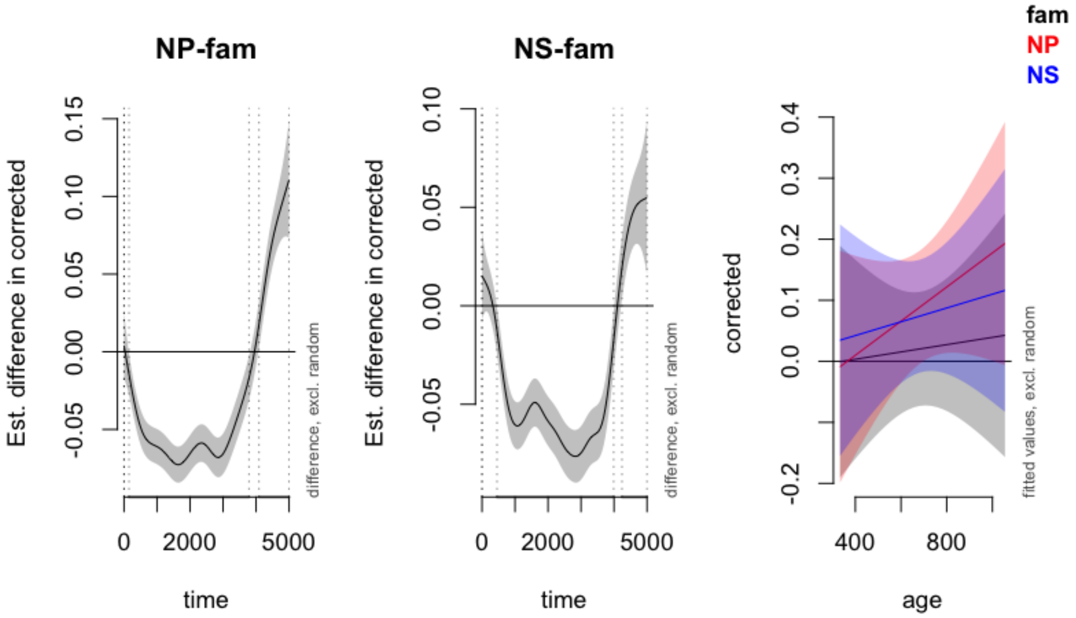


FIGURE 4.13: Estimated differences with point-wise 95% confidence intervals (M1, left panels) and estimated effect (M2, right panel). Left panels show the estimated difference between variations in pupil diameter in novel phonological (NP) and novel stress (NS) labels compared to familiar labels (fam) across time, as predicted in Model 1. Right panel shows the estimated effect of changes in pupil diameter across conditions (NP, NS, fam) in the test phase time window (4s to 5s) across age, as predicted by Model 2.

The results of this model (M1) are plotted in Figure 4.13, left. Specifically, the estimated differences between familiar and novel trials across time are plotted with point-wise 95% confidence intervals. As can be seen from these plots, the summary statistics of Model 1 indicate that the difference among the curves is significantly different from zero for both the novel phonological (b=-0.0271, SE= 0.0024, t=-10.923, *p*<0.001) and the novel stress (b=-0.0328, SE= 0.0024, t=-13.344, *p*<0.001) conditions. Specifically, the time window in which increases in pupil diameter toward novel labels were registered corresponds to the test phase of the trial (4s to 5s; see Figure 4.13). Lastly, the estimated difference is higher for novel phonological labels than novel stress ones (Est. difference NP: 0.15, Est. difference NS: 0.10).

Data from the test phase were then modeled by adding condition (i.e., trial type) per age and time as continuous predictors with participants and trials as random effects (M2). The resulting structure of Model 2 is described as follows and results are plotted in Figure 4.13, right. As can be seen from the plot, the summary statistics of Model 2 indicate a significant increase in pupil diameter with respect to familiar ones with age for novel phonological labels (b=-0.0873, SE= 0.0250, t=-3.487, *p*<0.001) but not for novel stress labels (b=0.0051, SE= 0.0243, t=0.209, *p*=0.834).

M2: *Pupil diameter   trialtype*age + s(time, by=trialtype, k=20) + s(time, id, bs='fs', m=1) + s(time, trial, bs='fs', m=1)*

Therefore, results from the linguistic task show that infants discriminate deviant phonological and prosodic features in recently familiarized label-object pairs, as signaled by significant increases in pupil diameter. Furthermore, phonological features seem to have a more detailed representation compared to prosodic ones and this trend seems to grow with age.

Link with rhythm

The next step was to explore the link between linguistic and rhythmic abilities. Performance scores in the rhythmic task were first computed from the ratio of averaged pupil diameter in the novel over total trials in the test phase [novel / (novel + familiar)], per participant. Variations in pupil diameter in the linguistic task were then modeled by means of GLM models (Nieuwenhuis et al., 2012; Fox, 2015) following a hierarchical step-wise model comparison. Fixed effects were gradually added to the null model (with random intercept per participant) and included the scores in the rhythmic task and age as continuous predictors and condition (i.e., trial type; fam, NP, NS) as a categorical predictor. As shown in Table 3, data were best explained by Model 4 (Wagenmakers and Farrell, 2004, lower dAIC and higher AICw). Specifically, Model 4 accounted for the interaction between rhythmic scores and condition, and the additive effect of age (Figure 4.14). Specifically, Model 4 predicted a significant interaction between the rhythmic scores and the pupil diameter in the novel phonological (b=0.2382, SE= 0.0233, t=10.208, $p<0.001$) and novel stress (b=0.1897, SE=0.0217, t=8.724, $p<0.001$) conditions (see Figure 4.14).

Increases in pupil diameter toward deviant phonological and prosodic are predicted by higher rhythmic scores. It is here important to specify that rhythmic scores were computed as the ratio of the pupil increases toward novel test trials [novel / (novel + familiar)]. Therefore, values above 0.5 indicate an increase for the novel stimulus while values below 0.5 indicate an increase for the familiar one. This means that infants who show an increase in pupil diameter for the familiar stimulus in the language task also showed an increase for the familiar stimulus in the rhythmic task, and vice versa.

TABLE 4.4: GLMM comparison for rhythm on language

Models	Deviance	dAIC	AICw
M.0 Pupil diameter ∼(1 | id)	1077	6.75	0.03
M.1 Pupil diameter ∼rhythm + (1 | id)	1161	92.05	0
M.2 Pupil diameter ∼rhythm + trial type + (1 | id)	1161	96.12	0
M.3 Pupil diameter ∼rhythm + trial type + age + (1 | id)	1175	112.52	0
M.4 Pupil diameter ∼rhythm * trial type + age + (1 | id)	1058	0.00	0.97
M.5 Pupil diameter ∼rhythm * trial type * age + (1 | id)	1078	29.66	0

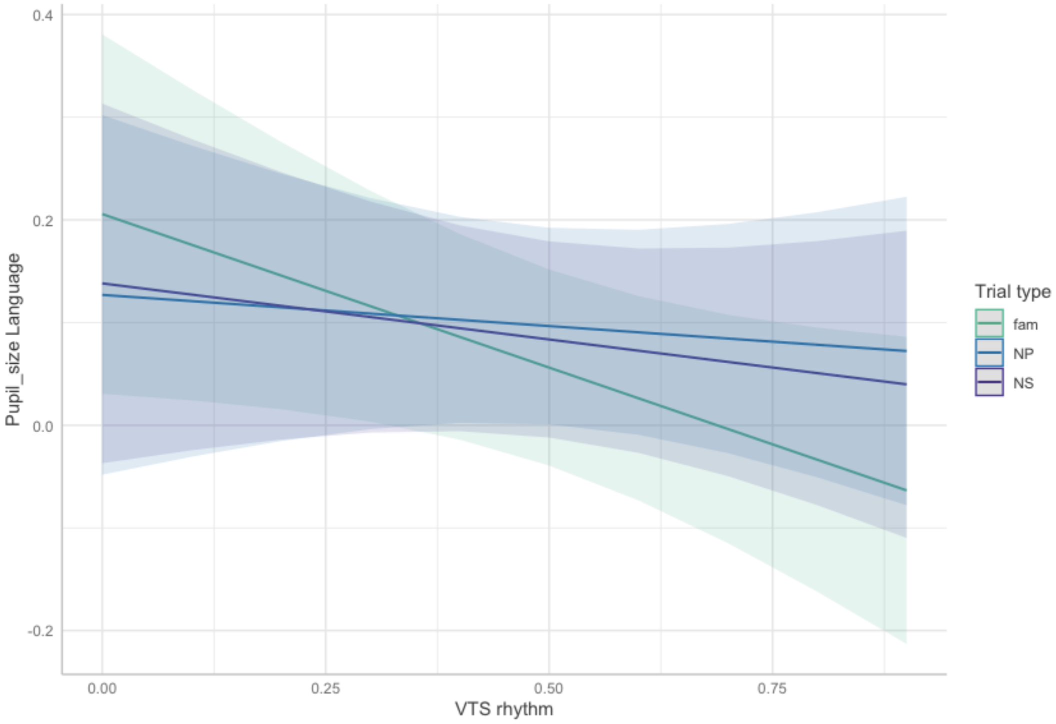


FIGURE 4.14: Effect plots of the best model (M4). Predicted relationship between variations in pupil size across conditions (i.e., trial type) in the linguistic task (i.e., familiar, novel phonological, and novel stress) and scores in the rhythmic task.

Since variations in pupil diameter were shown to be modulated by age in both tasks (see Figure 4.1.3 right and Figure 4.13), the next step was to explore whether age contributed to explaining the relationship observed between the rhythmic and language tasks observed in the results. Changes in the link between linguistic and rhythmic abilities across ages were explored by means of a GAM model. Percentages of averaged increase in pupil diameter for novel stimuli were computed for the two deviant conditions in the linguistic task with the same ratio used in the rhythmic one [novel / (novel + familiar)]. Then, individual scores across tasks (i.e., rhythmic, phonological, and prosodic) were taken as dependent variables and the interaction between tasks and age as a continuous predictor was modeled with the following structure. The resulting estimated effects are shown in Figure 4.15.

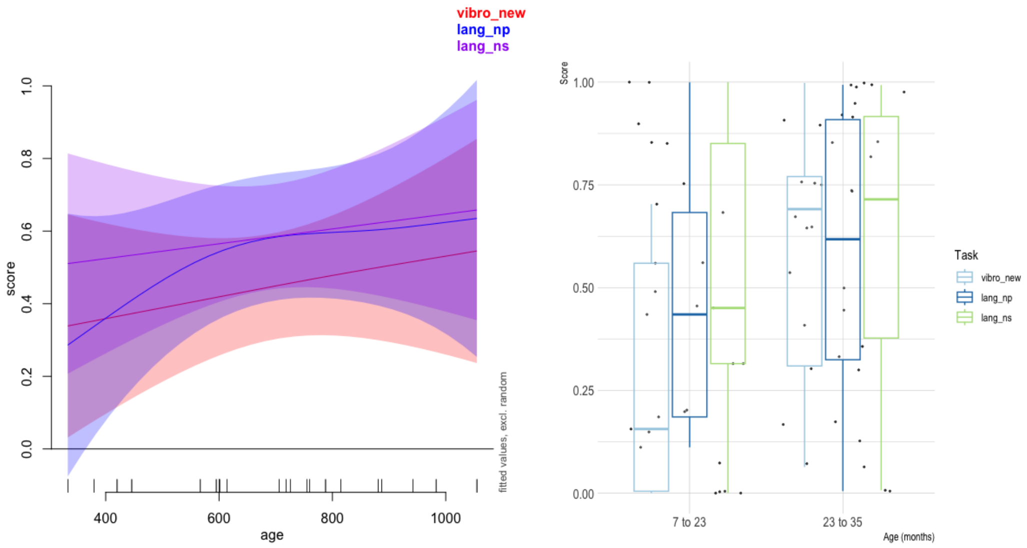


FIGURE 4.15: Estimated effect (Model 3) and descriptive statistics of task scores across age. Performance scores (y-axis) per task (i.e., rhythm as vibro new; phonological as lang no; and prosodic as lang ns) across age (x-axis). Estimated effect from Model 3 (left panel) and descriptive statistics with individual performance (right panel).

M3: *Scores task + s(age,task, k=20) + s(id, bs='fs', m=1)*

As the plot shows, scores grow with age with no significant difference between tasks (NP: b=0.0933, SE=0.0918, t=1.016, p=0.315; NS: b=0.1406, SE=0.0917, t=1.533, p=0.132). Paralleling the findings from the linguistic task analysis (see Figure 4.13, right), scores in the phonological task show a rapid increase compared to prosodic and rhythmic tasks, even if differences are not significant. Results will be discussed in the next section.

4.2.4 Interim discussion

This Study investigated the emerging linguistic abilities of 7- to 35-months-old infants. Specifically, phonological and prosodic abilities were assessed through a familiarization-test paradigm for object-label pairs. In this paradigm, infants were familiarized with novel object-label pairs; then, they were presented with the original pairs (i.e., familiar condition) or with labels mispronounced for phonological (i.e., novel phonological condition) or prosodic (i.e., novel stress condition) features. Increases in pupil diameter toward corresponding visual referents (i.e., objects) were collected through the eye-tracker and analyzed as the dependent variable. Data were then analyzed by means of a GAM model accounting for condition (i.e., fam, NP, NS) as a fixed, categorical effect and age as a continuous predictor of pupil changes across time, with random intercepts for participants and items. Results show a significant increase in the pupil diameter occurring in the test phase (i.e., the time window going from 4 s to 5 s, in which labels were pronounced based on condition). Moreover, the model predicted an increased estimated difference between familiar and novel phonological items (i.e., 0.15) with respect to familiar and novel stress items (i.e., 0.10). After this, another model was run to investigate the changes occurring in the test phase across ages (Figure 4.13, right). Results show that pupil diameter increases for test over familiar conditions in general and for changes occurring on phonological features in particular.

Overall, these results suggest a fine lexical representation across ages. Specifically, object-label pairs seem to be encoded with detailed phonological and prosodic features. This result is in line with previous findings reporting mispronunciation effects on word recognition (Durrant, 2017; Swingley and Aslin, 2000; Yoshida et al., 2009; Delle Luche et al., 2015). However, most of these studies consisted in a word recognition task, where two visual objects were presented on a screen (i.e., a target and a distractor) and labels could be presented in corrected or mispronounced form (Hohle et al., 2006; Mani and Plunkett, 2011; Swingley and Aslin, 2000; White and Morgan, 2008; Delle Luche et al., 2015). Looks toward the target object have been found to be slower and less accurate when labels are mispronounced. This led the authors to the conclusion that early lexical representations are well specified, with mispronunciations impacting word recognition (Mani and Plunkett, 2011; White et al., 2013). By contrast, in the present Study a new paradigm was developed allowing us to better specify the effect of segmental and suprasegmental features on object-label encoding, while minimizing the impacts of other incidental factors. Indeed, the presence of a distractor, and the use of real words as behavioural measures represent potential sources of variability that can be reduced by increasing the methodological strength (Klingner, 2010).

Specifically, in the present Study infants were presented with novel object-label pairs through a familiarization-test paradigm; furthermore, labels were pseudo-words and mispronunciations occurred right after the familiarization phase, minimizing possible external interference of time lag. Indeed, every trial consisted of a familiarization and a test phase (Russo et al., 2021; Calignano et al., 2021a). Specifically, this allowed us to test up to three novel object-label pairs, improving the generalizability of our observations. Lastly, the dependent variable of this study was the change in pupil diameter occurring in correspondence to the test phase, which was within the time window (4s to 5s) in which labels were pronounced based on the experimental condition. This led to a narrowed response since changes in pupil diameter are a well-established physiological index of cognitive effort (Karatekin, 2007).

Furthermore, the statistical approach beyond this Study – employing GAM and GLM models – allowed for modeling changes in pupil diameter across the time dimension while dealing with individual variance for participant and item as well as exploring the effect of age as a continuous predictor. By doing so, it was possible to show that: i) infants are sensitive to changes in phonological and prosodic features occurring in recently familiarized object-label pairs; ii) that phonological features are more detailed than prosodic ones, and iii) they grow in specificity from 7 to 35 months. This result replicates, with an experimental paradigm and a dependent measure different from the literature, that early lexical encoding is highly specified in infancy (Durrant, 2017; Swingley and Aslin, 2000; Yoshida et al., 2009; Delle Luche et al., 2015).

Moreover, this result brings new insights into the cognitive mechanisms underlying this process and the developmental changes that occur across ages. Specifically, pupillometry measures are fine enough to show that infants are sensitive to the different degrees of mispronunciations spanning from the less detailed suprasegmental features (i.e., novel stress condition showed a different effect of 0.10) to the more specific segmental ones (i.e., novel phonological condition showed a different effect of 0.15). This result is in line with the recent evidence from Tamási et al. (2017, 2019) showing different degrees of response based on the number of altered features (i.e., manner of articulation, voicing, and place of articulation). The present Study further shows that alterations at the suprasegmental and segmental levels differently impact label recognition in infancy. Furthermore, infants are here shown to develop a higher response (in terms of pupil dilation to mismatching items) for phonological alterations with age.

Altogether, these findings can be interpreted as a sign of the infant shift from basic to fine linguistic abilities in lexical encoding. Specifically, prosodic cues are known to bootstrap the first steps of language acquisition in young infants; while fine phonological abilities are known to develop with age, with perceptual narrowing mechanisms aiding the processing of native phonological features as a function of experience (Kuhl, 2004; Gervain, 2018; Jusczyk et al., 1999). Crucially, previous studies investigated the effect of mispronunciations on word recognition; by contrast, pupillometry is known to better reflect cognitive effort. This means that a higher cognitive effort was required to process altered labels. This result is more indicative of the mismatch between the familiarized object-label pair and the mispronounced one rather than word recognition. Therefore, it can be inferred that labels varying in prosodic features require less cognitive effort to be processed than phonological ones and that phonological representations grow in specificity with age requiring more cognitive effort to be processed when a mismatch is detected.

This would be an adaptive strategy for infants since segmental information changes at a fast rate in speech and is highly informative in terms of conveyed meaning (Fernald et al., 2001, e.g., /bat/ and /cat/ have two different meanings driven by a single change in phoneme). By contrast, suprasegmental features like prosodic characteristics are known to vary between different speakers and emotional intonation (Selkirk, 1995). Accordingly, while young infants strongly rely on prosodic cues moving their first steps in language acquisition, the possibility to develop a greater degree of flexibility in prosodic features could characterize later steps of lexical representations (Seidl and Cristià, 2008). Therefore, this developmental interpretation might explain: i) the larger effect obtained in the novel phonological condition, and ii) the growth of this effect along with age. At the same time, even with a less robust effect and with no significant growth along with age, infants show an increase in pupil diameter in response to novel stress labels as well. Once the differences between the two manipulations (i.e., phonological and prosodic) have been explained in terms of developmental language-specific cue weighting, it can be clarified that changes in prosody remain informative even if the infant cognitive system treats with more flexibility those changes containing the impact on word recognition.

Our analysis further continued by exploring the link between the linguistic and rhythmic abilities registered in Study one and Study two. Overall, the performed comparison approach to GLM models allowed for identifying the best model as the one accounting for the interaction of rhythmic score and condition, with the additive effect of age, on pupil changes occurring in the linguistic task. This means that changes in pupil diameter during the linguistic task were best explained by the experimental manipulation (i.e., fam, NP, and NS conditions) interacting with individual scores in the rhythmic task, considering the effect of age as well. It can be recalled that individual scores in the rhythmic task result from the percentage of averaged pupil diameter for novel stimuli in the test phase, per participant. It ranges from 0 (for infants showing no increase in pupil diameter for novel stimuli) to 1 (for infants showing no increase in pupil diameter for familiar stimuli). Therefore, as the plot shows (Figure 4.14), infants showing lower values of pupil increase for novel stimuli (i.e., higher increase for familiar stimuli) in the rhythmic task also show a higher increase in pupil diameter toward familiar stimuli in the linguistic task. On the contrary, infants showing an increase in pupil size for novel stimuli in the rhythmic task also show an increase in pupil size for novel stimuli in the linguistic task. This result tells us that: i) there is a link between performances in the rhythmic and linguistic tasks; ii) this link has a composite pattern rather than an absolute value. This last point is particularly true considering that every individual behaves differently, especially across development since even a few months can make the difference. Regarding the link between the two tasks, according to the performed statistics (see Section 4.2.3), there is a significant interaction between rhythmic scores and performances in both the novel phonological and novel prosodic conditions.

Therefore, this result suggests that the way in which infants react to unpredicted changes occurring in the underlying meter of musical rhythms might be related to the way in which they react to unpredicted changes in the phonological and prosodic features of linguistic stimuli. Furthermore, stimuli in the rhythmic task were perceived as VTS stimuli while labels in the linguistic task were presented only as auditory. This means that an overlapping set of basic rhythmic abilities might serve the processing of a vast range of signals across domains (i.e., from music to language) but also across sensory modalities (i.e., from touch to hearing). This result is particularly informative considering the growing literature on the link between basic rhythmic skills in language acquisition (Fiveash et al., 2021, see Chapter 3) and cross-sensory perception (Karam et al., 2009b, see Chapter 3). Indeed, the first framework and theories on rhythm in language development were recently born and called for a detailed investigation into the link between cross-domain and cross-sensory abilities underlying rhythm processing in musical and linguistic domains. Accordingly, the present Study brings significant new insights into the extent to which rhythm processing might be considered a general cognitive ability serving the processing of different temporal signals across modalities.

Regarding the limitations of this hypothesis, it could be argued that performances in the two tasks are related because of the same infants participated in the two studies and the same measure was collected in the two tasks. Therefore, the emerging relationship between results could be better interpreted as a general cognitive profile stable for individuals. Given that physiological measures allow for a clearer picture of the underlying processes than behavioural measures, but data remain open to interpretations, this hypothesis cannot be completely rejected. However, even if what has been observed here was the actual cognitive profile of each infant participating in the studies, this would exactly be what was aimed to be collected in this work. Specifically, the high similarity between tasks in terms of experimental design, procedure, and collected measures is a good rather than a problematic practice when comparing tasks in psychology. Furthermore, identifying possible ways to compare rhythmic and linguistic measures is one of the main challenges in the research field (Ladányi et al., 2020; Fiveash et al., 2021).

Specifically, the present Study shows that – with methodological and statistical strength – individual variability can be modeled thus allowing for comparisons between different abilities. Moreover, since the evidence in favor of the link between rhythmic and linguistic abilities keeps growing, the next step nowadays is to identify the cognitive and neural mechanisms underlying the overlap between these two skills. To this aim, the present Study brought evidence on the cognitive mechanisms underlying the processing of rhythms and language stimuli. Specifically, it has been shown that the cognitive effort sustaining the processing of deviant underlying meters in rhythm encoding is related to the cognitive effort needed to process recently familiarized labels altered for phonological and prosodic features. Therefore, the present result might contribute to advancing our knowledge by bringing new evidence on the shared mechanisms underlying cross-domain and cross-sensory processing of rhythm.

Furthermore, as mentioned a few lines before, the link between rhythm and language abilities observed here shows a composite pattern rather than an absolute value. This means that infants seem to behave similarly across the two tasks but they also show individual differences. Specifically, the data shows that there are infants processing novel stimuli with a higher cognitive effort and infants who require more cognitive load to process familiar rather than novel stimuli. This result can be explained not only by considering individual differences (and the aforementioned individual cognitive profile) but also by stressing the fact that a large range of developmental stages was here included, going from 7 to 35 months of life. Moreover, the age effect was also reported at the single study level in both tasks as expected since developmental changes were one of the main research questions of this work. Specifically, the aim was here to explore rather than avoid variance across development, accounting for age as a continuous variable rather than zoom into a specific age window. To this aim, the last GAM model was run accounting for the effect of task (i.e., rhythmic, phonological, and prosodic) as a categorical predictor on performances across the age continuum. To run this model, performance scores for the linguistic task were performed with the same ratio as the rhythmic.

Therefore, individual scores in the tasks were the dependent variable of the model. As shown in the result section, performances show a slight growth along with age, with no significant differences between tasks. Again, it must be recalled that the term score here indicates the percentage of increase in pupil diameter toward novel stimuli. Therefore, it can be said that infants increase their pupillary response similarly toward novel rhythmic and linguistic stimuli with age. A greater increase, but not significant, seems to occur for phonological abilities, paralleling results at the single study level. This exploratory result might be first interpreted as a possible track to explain the observed in the previous analysis and, secondly, as a developmental trajectory in the cognitive abilities of infants.

Specifically, as their experiences with native language and music patterns grow, the cognitive ability to extract an invariant from rhythmic and linguistic signals grows as well, as demonstrated by the increased cognitive effort required to process deviant forms. Moreover, this seems to be a general ability serving different perceptual and cognitive domains since it seems to be a link across sensory modalities (i.e., touch to hearing) and cognitive domains (i.e., music and language). This link might be further explained by the similarities between these signals. Indeed, music and language are complex structures varying in time, hierarchically organized, and perceived through vibratory patterns. The vibratory nature of both signals might indeed further explain the link between the two tasks, especially considering the core role of VTS modality in coding rhythm. Lastly, the present Study contributed to defining rhythm perception as a multidimensional ability since perceptual channels in the human body are strictly connected as well as the way in which they are processed in the brain. Therefore, the main insight here is that infants strongly rely on this rich structural information to make sense of the world. Specifically, rhythm might allow infants to extract the structure underlying the incoming inputs to better direct their attentional resources and maximize their processing abilities. According to the present Study, this seems to happen with both musical and linguistic signals, as well as with tactile and auditory input. Furthermore, this ability seems to increase with age representing a potential developmental scaffold for cognitive development and language acquisition in particular.

4.2.5 Conclusions to Study two

Now that results have been discussed in terms of visual inspection and statistical significance, it may be time to recall the initial question motivating this Study and see how the observed results might provide an answer. Specifically, this Study aimed at investigating the emerging linguistic abilities of infants from 7 to 35 months through a discrimination task for novel object-label pairs. Specifically, prosodic and phonological abilities have been measured by investigating how infants react to mispronunciations, occurring at the stress or the phonological level, in recently familiarized labels. Furthermore, this Study aimed at testing whether the linguistic abilities here collected do relate or not to the rhythmic skills measured in Study one, and how this relationship might change across age. Based on results, this Study might provide useful information to answer these questions. Specifically, early lexical representations seem to be highly specified in the infant mind and this can be inferred by the increment in the cognitive effort (signalled by pupil dilation) needed to process mismatching labels varying for phonological and prosodic features.

Changes in phonological features in particular impacted label processing, requiring more cognitive effort to be processed and this trend seems to increase with age. Furthermore, performance scores in the rhythmic task significantly contribute to explain results in the linguistic task, providing evidence in favour of a link between the two domains. Specifically, rhythmic abilities might be informative about linguistic skills. Lastly, the relationship between rhythmic and linguistic skills seem to proceed with age, with increasing performances across tasks. Specifically, what seems to grow with age is the amount of attentional resources invested in the processing of novel stimuli, being these musical or linguistic signals. This could be a cognitive boost to sustain learning processing (Yu et al., 2019). Therefore, the present Study brings new evidence into the perceptual and cognitive mechanisms underlying the processing of rhythm and language stimuli across infancy. Furthermore, it contributes to identifying the extent to which the overlapping mechanisms serving the process of rhythm in music and language also encompass different sensory modalities. Indeed, even if further research is needed to replicate and extend the present findings, this Study is the first to our knowledge to show a link between the processing of musical rhythms perceived only via VTS inputs and phonological and prosodic abilities perceived through auditory input in infancy.

Therefore, the present findings might contribute to inform the future theoretical models of cognitive development as well as the next steps of cognitive research, with possible implications for health and education. To this aim, the next step of this work was to test whether the well-known benefits originating from musical rhythm might be transferred not only from one domain to another (i.e., music to language) but also across sensory modalities (i.e., tactile to auditory). If so, then important implications for the infant vibrotactile system might arise in training early cognitive abilities supporting language acquisition. Study three aims at providing preliminary evidence on this topic, as discussed in the next section.

4.3 Study three: a pilot VTS rhythmic priming for processing language

Based on previous studies showing that infants are sensitive to rhythmic structures perceived with the VTS sensory modality (first study) and that this ability correlates with emergent language abilities (second study), this study will directly test the benefit on processing language when primed by a matching, short-term VTS rhythmic exposure. First, infants will be presented with a short rhythmic prime by means of the Infant vibrotactile system described in Study one. Then, VTS rhythmic primes will be followed by matching or mismatching linguistic sequences presented in the auditory sensory modality. Specifically, a rhythmic prime with a double or a triple meter will be followed by a disyllabic (i.e., matching for double and mismatching for triple meter) or a trisyllabic (i.e., mismatching for double and matching for triple meter) linguistic sequence. Changes in pupil diameter toward contingent visual stimuli to matching or mismatching trials will be collected as an index of cognitive effort (Beatty et al., 2000; Arias-Trejo et al., 2022). Based on the literature reviewed in Chapter 3, the possibility to benefit from a rhythmic prime in processing linguistic sequences is expected in this Study. Accordingly, rhythmic priming tasks have been largely applied to investigate the possibility to transfer benefits from short-term exposures to musical rhythms across domains (Cason and Schön, 2012; Cason et al., 2015a; Fiveash et al., 2020; Gould et al., 2017; Chern et al., 2018). Moreover, the rhythmic priming effect was found to occur also across different sensory modalities, including auditory to visual (Fotidzis et al., 2018) and visual to auditory signals (Liu et al., 2012; Gould et al., 2018). However, very few studies investigated the cross-modal influence of haptic inputs (Schneider et al., 2011), especially considering sensorimotor and vestibular modalities which have been investigated only as a cue to spatial orientation (Lindeman et al., 2003).

Therefore, the present Study was designed to test whether the Infant vibrotactile system could serve as a possible application for transferring the documented effects of rhythm on language processing through short-term exposure to VTS rhythmic primes. Specifically, this third Study was directly motivated by the encouraging results provided in Studies one and two. Accordingly, given the documented efficacy of the Infant vibrotactile system in testing VTS rhythmic abilities in infancy (see Study 4.1) and its link with emerging linguistic skills (see Study 4.2), it has been decided to further test the potentialities of this work in an applied context. For this reason, rather than leaving open the possibility to investigate the direct effects of VTS rhythms on language processing in future studies, it has been decided to design and start to implement this paradigm as a pilot to provide a concrete step for directing my future research. For this reason, a pilot of this Study is introduced here with preliminary data presented and discussed based on their potential implications for future research. As for Studies one and two, data collection was made possible thanks to the contribution of 2 kindergartens as part of the Cooperative Now Project. Design and preliminary results were discussed with researchers from University of Padua and abroad during meetings, research stays, and project presentations.

4.3.1 Aim and research questions

The present Study aims at investigating the possibility to benefit from a VTS rhythmic prime in processing language. That is, after investigating the cross-domain and cross-sensory links between VTS rhythms and linguistic abilities in infancy, the present project aims at advancing the current knowledge into the potential impacts of nowadays research on music and language development. Specifically, one relevant aspect of this research line is the idea that cross-domain effects in one domain might occur after training the other. This has been traced back to the similar mechanisms underlying music and speech processing, with temporal regularities (governing both signals) allowing for expectations and prediction (McAuley, 2010; Arvaniti, 2012).

Therefore, rhythmic priming is hypothesized to provide a scaffold to organize the following linguistic signals (Flaten et al., 2022; Brown, 2012; Benasich and Tallal, 2002). Therefore, the present Study aims at investigating the cross-domain dynamics between rhythm and language by focusing on the effects generated by short-term exposure to musical rhythms. Moreover, the present Study aims at advancing our knowledge on the cross-sensory effects of rhythm on language processing (Fotidzis et al., 2018; Liu et al., 2012; Gould et al., 2018, see Chapter 3) by providing rhythmic priming under the VTS modality. Indeed, the core involvement of sensorimotor processing in rhythm perception is now fully recognized in the literature (Nayak et al., 2022; Stephan et al., 2016) with infants in particular being found to rely on sensorimotor information to organize rhythmic stimuli (Phillips-Silver and Trainor, 2005) and audio-motor training enhancing rhythmic effects on language processing (Cason et al., 2015a, see Chapter 2 and 3). However, the potentialities of haptic sensory modality and VTS inputs in particular are still significantly underestimated (Bremner and Spence, 2017, see Chapter 2).

Therefore, the main research question of this Study is whether the cross-domain (i.e., musical to language) and cross-sensory (i.e., tactile to auditory) associations between rhythm and language do sustain the possibility to rely on one signal and modality to process the other. This last step will therefore contribute to clarifying the scaffolding role of rhythm in language development by providing new insights into the extent of cross-domain and cross-sensory transfers of benefits. This will be made possible thanks to the technological and methodological advantages applied in this project, including: i) the Infant vibrotactile system originally developed in Study 4.1; ii) the implementation of pupil data collection and gaze-triggered paradigms to study the underlying mechanisms of early language processing, as illustrated in Study 4.2; and iii) the employment of generalized (additive and linear) mixed models and model comparison to test for significance while modeling data as continuous measures (maximizing the number of observations per trial and participants as well as the possibility to specifying the data distributions) and accounting for inter-individual variability. Therefore, the research method will be illustrated as follows and preliminary results will be then presented and discussed based on potential implications for language development.

4.3.2 Method

Participants

Ten infants (6 females, M age = 632 days, SD = 306,68) participated in this pilot Study, recruited from kindergartens from Padua (IT) and adjacent municipalities as part of the Cooperative Now Project. The inclusion criteria for all participants were: to be in good health, to have no sensory or neurological disorders or any familiar language disorder, and to be native Italian speakers. The research protocol was approved by the Ethics Committee of the Human Inspired Technology research center at the University of Padua (protocol number: 202066R1). The research was conducted in accordance with the principles of the Declaration of Helsinki. Caregivers signed a consent form before taking part in the experiment while verbal consent was asked to participants before starting the experimental section.

Stimuli

Stimuli consisted of VTS rhythmic primes and auditory linguistic sequences presented in contingency with a visual attentional getter. The primes were two different rhythms, one displaying a double meter and the other one displaying a ternary meter (Cason et al., 2015a, Figure 4.16). As in Study 4.1, the VTS rhythms were originally synthesized as audio tracks with the library SoCal drum sound in GarageBand and then presented as VTS stimuli by means of the vibrotactile system. Rhythms were followed by disyllabic or trisyllabic linguistic sequences, matching or not the prime metric structure (see Figure 4.16). Linguistic sequences were presented auditory. As in rhythmic so in linguistic sequence, stress was always placed on the first event (i.e., notes in rhythms and syllables in language; Figure 4.16, left). The binary prime had a duration of 4 seconds while the ternary prime of 6 seconds; disyllabic stimuli had a duration of 1 second while trisyllabic stimuli of 1.5 seconds. In each trial, a prime was followed by a syllabic structure. For the entire duration of each trial, a contingent visual stimulus was presented and consisted of a colourful checkerboard on a neutral background (Figure 4.16, right). The image areas corresponded precisely to the areas of interest of the eye-tracker (AOI), which measured 10 × 10 cm (9.554 deg).

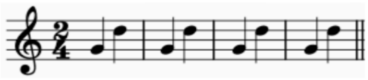
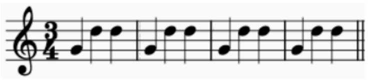
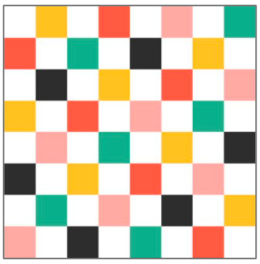
Prime	Match	Mismatch
	/ **Ga** ga /	/**Ga** ga ga /
	/ **Co** co co /	/ **Co** co /



FIGURE 4.16: Stimuli from Study three. On the left, the binary (i.e., 2/4) and ternary (i.e., 3/4) priming followed by matching or mismatching syllabic sequences. On the right, the colourful checkerboard shown as visual stimulus.

Apparatus

The experiment was programmed and presented through Open Sesame software 3.1 (Mathôt et al., 2012) running on a computer laptop (Acer travel mate 5772 g). Rhythmic primes were presented under the VTS modality through the Infant vibrotactile system. This system, extensively described in Study one (see Section 4.2 of this Chapter), is a music transduction device specifically designed for this project and based on the Model Human Cochlea (MHC) developed by Karam et al. (2009b). Specifically, it translates auditory information into vibrotactile stimuli using a tactile display. Music is displayed through voice coils facilitating the direct translation of auditory information onto multiple discrete channels projected toward different portions of the back. The syllabic sequences were auditory and presented through two loudspeakers placed on both sides of the monitor. Visual stimuli were displayed on a 27-inch monitor (Philips 300 × 300) while a remote, infrared eye-tracking camera (Tobii X2-60 Eye-Tracker) placed directly below the screen, 60 cm away from the participant, recorded the eye movements using bright-pupil technology at a sampling frequency of 60 Hz.

Experimental design and procedure

To directly investigate the cross-modal effects of VTS rhythms on language processing, this Study implemented a gaze-triggered, rhythmic priming paradigm (Figure 4.17). Variations in pupil dilation were continuously recorded during the experiment, eliciting the trial presentation.

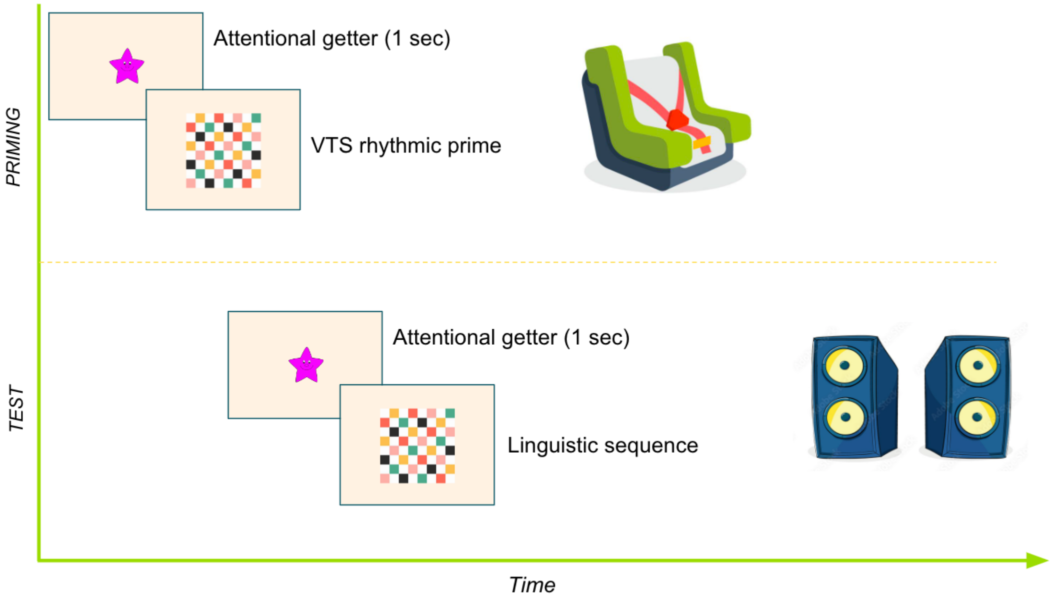


FIGURE 4.17: The cross-modal, rhythmic priming paradigm from Study 3. Each trial consisted of a priming and a test phase. The rhythmic priming was administered under the VTS modality through the Infant vibrotactile system placed on the infant seat; then, the linguistic stimulus was auditory presented. Contingent visual stimuli (i.e., colourful checkerboard and a cartoon attentional getter) were presented to attract the infant's attention allowing the eye-tracker to collect pupil data.

At the beginning of each experimental section, a calibration was run. Participants were presented with small, cartoon images appearing on the screen at five locations (i.e., top-left, top-right, center, bottom-left, and bottom-right) together with a piece of cheerful music to capture their attention. Once the calibration was successfully completed, the experiment started. The gaze-triggered procedure implemented in this Study ensured that each experimental trial only began once the infant's gaze was detected as continuously directed for more than 500ms toward the AOI by the eye-tracker and it stopped when infants looked away for more than 2 seconds. Between trials, an attentional getter was presented on the screen to attract the infant's gaze toward the monitor. Each trial consisted of a prime and a linguistic sequence. For each trial, once the gaze of participants was detected as inside the AOI and continuously registered for 500 ms, a VTS rhythmic prime was first presented.

After the participant's gaze was redirected toward the screen by the attentional getter, a silent time window of 500ms allowed the eye-tracker to collect baseline data on the pupil diameter prior to the linguistic sequence occurred. Linguistic sequences were then auditory presented and, together with the following 1.5s of silence, this phase was taken as a test window in which differences in pupil diameter were expected (Mathôt and Vilotijević, 2022). Therefore, each trial had a minimum of 6.5s and a maximum of 10.5s of duration, based on the combination of prime and syllabic sequence.

Specifically, this Study implemented a 2x2 factorial design where the first factor was the prime meter (i.e., binary/ternary) and the second factor was the syllabic sequence (i.e., matching/mismatching). Therefore, binary matching trials were shorter in duration (i.e., 6.5s) than ternary mismatching trials (i.e., 10.5s); however, only the linguistic presentation was explored and the following time window of 1.5s which was equal to all trials regardless of their total duration. Each combination was presented six times, resulting in 24 trials. The presentation order was counterbalanced with 10 participants per combination. Trials were divided into two blocks, one displaying the matching and mismatching trials of the first meter and the other one displaying the matching and mismatching trials of the second meter. Between blocks, a short cartoon video was presented as a break before the second block presentation. During the experimental section, infants were seated on a highchair positioned 60 cm away from the monitor and were tested individually in a quiet room of their kindergarten. Kindergarten educators familiar with the infants assisted in the experimental session to comfort them in case of crying or fussiness. Semi-darkness constant luminance was guaranteed by a lamp positioned 1m behind the participant.

Statistical analyses

Only data from participants who completed at least one of the two experimental blocks were analyzed. All participants from this pilot were presented with the binary block first. Since not all participants succeeded in completing the ternary block, data from the binary block only were included in this preliminary analysis. Variations in pupil diameter registered after matching or mismatching stimuli were presented were taken as the dependent variable. Specifically, changes in pupil size under constant luminance were taken as a measure of cognitive processing during stimuli presentation (Beatty et al., 2000; Einhäuser, 2017). Pre-processing steps (i.e., transforming the raw data into a format for statistical analysis and visualization) were performed following Hepach and Westermann (2016) procedure and (Mathôt and Vilotijević, 2022) guidelines (Figure 4.4). Since the pre-processing procedure here applied is the same as in Study one and Study two, please refer to section 4.1.2 of this Chapter for an extended description of each step. Pre-processed data will be visually inspected and analyzed by means of Generalized Mixed-Effects models (Nieuwenhuis et al., 2012; Fox, 2015). Fixed effects (i.e., block: binary/ternary and condition: matching/mismatching) will be gradually added to the null model (only including the random effect of participant) and then the model comparison will be performed using the likelihood ratio test (LTR), Akaike Information Criterion (Akaike, 1974, AIC), and AIC weight as indexes of the goodness of fit. Since for this pilot data from binary block only were included, preliminary analyses will consist of visual inspection and a simple GLM model accounting for the effect of condition (i.e., matching/mismatching), as described below.

4.3.3 Preliminary Results

Descriptive statistics (Figure 4.18) show a slight increase in pupil diameter during the test phase for mismatching over matching trials. Specifically, the data distribution, shown by violin (Figure 4.18, left) and density (Figure 4.18, upright plots), is shifted over positive values for pupil diameter in mismatching trials compared to data from the matching condition.

Even though a small sample size (N=10) has been tested for this pilot Study, a feasible number of observations per participant (Figure 4.18, bottom right) is registered in pupillary data. To better explore variations in pupil diameter over time, baseline corrected pupillary data were plotted in the time dimension applying a locally estimated scatter-plot smoothing (Cleveland and Devlin, 1988, loess) that is a non-parametric regression method combining multiple regression models in a k-nearest neighbour, based meta-model (Jacoby, 2000). The loess method was specified in the geom_smooth function of the ggplot2 package (Wickham, 2016) in R software (R Core Team, 2022).

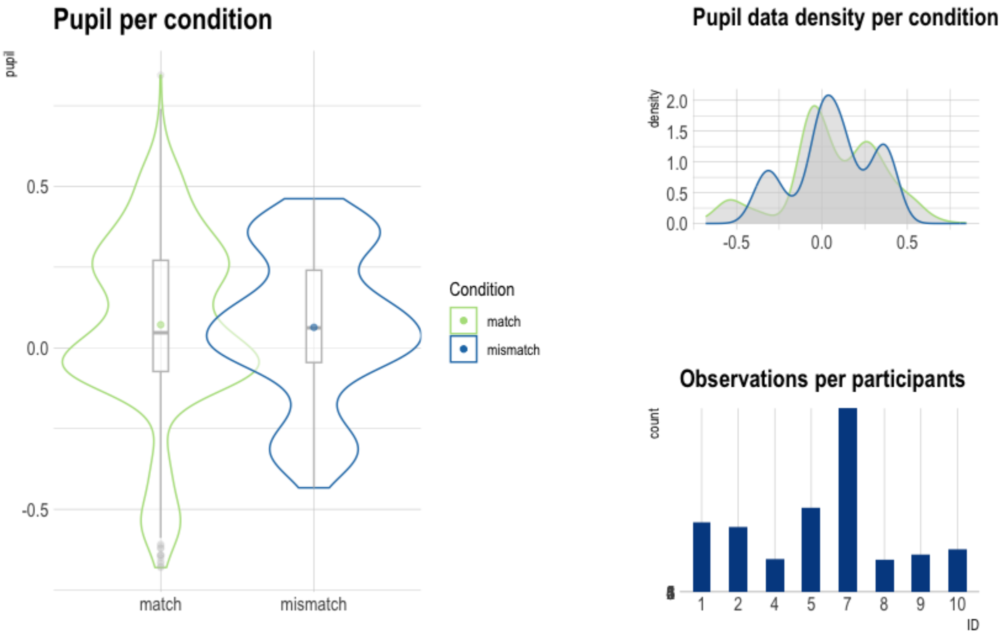


FIGURE 4.18: Descriptive statistics from pilot Study 3. Violin plots (left panel) of baseline corrected pupillary data per condition (i.e., matching / mismatching); density plot (upright panel) showing the distribution of baseline corrected pupillary data per condition (i.e., matching / mismatching). Number of observations (bottom right panel) per participant.

As Figure 4.19 shows, pupil diameter increases after mismatching stimuli were presented in the test phase compared to matching test trials (Figure 4.19, left). Pupil data were modeled with the GLM models shown in Table 4.5. Starting from the null model (M0) accounting for the random effect of inter-individual variance only, the fixed effect of condition (i.e., matching/mismatching; M1) was added. Model comparison through Akaike information criterion indicated M1 as the best model explaining the data distribution. Specifically, M1 predicted a significant increase in pupil diameter in the mismatching over matching condition (b=0.1002, SE=0.0955, t=1.049, $p<0.001$). These preliminary results will be discussed in the next section.

TABLE 4.5: GLMM comparison for pupillary data in Study 3

Models	Deviance	dAIC	AICw
M.0 Pupil diameter ∼(1 ∣ id)	-718.47	2.54	0.22
M.1 Pupil diameter ∼condition + (1 ∣ id)	-723.02	0.00	0.78

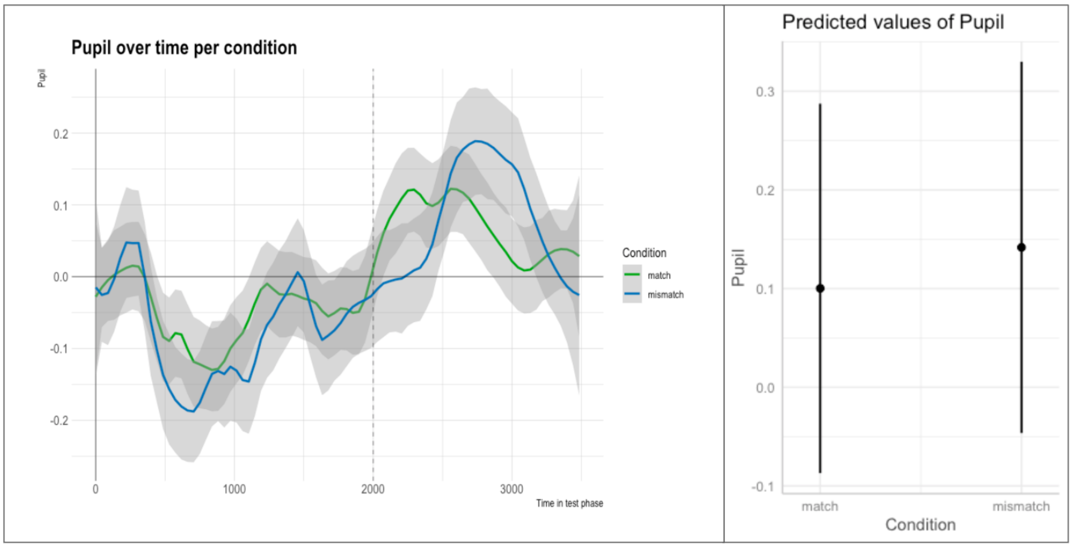


FIGURE 4.19: Pupil data from pilot Study 3. Time course of baseline corrected pupil data in the test phase per condition (left panel). The dotted line indicates the end of linguistic sequence presentation and the start of the silent window in which differences in pupil response were expected. Predicted values of baseline corrected pupil after matching/mismatching stimuli from M1.

4.3.4 Interim discussion

This pilot Study investigated the effects of a VTS rhythmic priming on auditory perceived linguistic sequence in infancy. Specifically, infants were presented with a rhythmic prime followed by a syllabic sequence matching or not the prime meter. Changes in pupil diameter following matching or mismatching stimuli were collected as an index of cognitive effort (Beatty et al., 2000; Einhäuser, 2017). Based on the previous studies showing that infants are sensitive to rhythmic structures perceived under VTS sensory modality (Study 4.1) and that this ability is related to emerging linguistic skills (Study 4.2), infants were expected to benefit from VTS rhythmic priming in processing following linguistic sequence perceived in the auditory modality. Specifically, an increase in pupil diameter was expected for mismatching trials signaling that the infants detected the deviance in the rhythmic structure of linguistic sequence requiring an increase in cognitive processing. Preliminary results from ten participants seemed to be in line with this hypothesis, showing an overall slight increase in pupil diameter in response to mismatching over matching stimuli (Figure 4.18).

Moreover, data were explored in a time course fashion showing that pupil diameter increased specifically after mismatching linguistic sequences were presented (Figure 4.19). All subjects participating in this pilot were assigned to the first order combination in the counterbalanced design, being tested with binary rhythmic priming first. Subjects completing at least this first block were included in this exploratory analysis. Therefore, the subtle delay in the mismatching response was probably due to the fact that mismatching stimuli in the binary block (i.e., syllabic sequences with a ternary meter) were 500ms longer in duration than matching ones. Accordingly, the peak in pupillary response to mismatching stimuli occurred around 500ms after the matching peak response. Data within the silent time window following stimuli presentation were modeled with GLM models and the model comparison based on Akaike information criterion showed that M1, accounting for the fixed effect of condition and random individual intercept, was the best model in explaining the data distribution. The plotted effect (Figure 4.19, right) showed a significant increase in pupil response in the mismatching condition.

However, the reported statistics are better representative of the tested sample rather than being extendable to the infant population given the small sample size (n=10). In fact, further research is needed to corroborate these preliminary findings. Specifically, testing infants in the ternary prime block (as planned in the counterbalanced design) is essential to disambiguate whether the observed increase in pupil diameter for mismatching trials is elicited by the priming effect or by the trisyllabic stimuli *per se*. In fact, infants growing up in Western cultures are known to show a bias for binary over ternary musical meters (Bergeson and Trehub, 2006). Accordingly, a binary bias was observed in the infant discrimination abilities for VTS rhythms in Study one. Therefore, a possible interpretation of these preliminary results could concern the processing bias for binary meters in infancy. That is, ternary sequences might have required a greater allocation of attentional resources *per se*. However, in processing language, infants have been found to correctly encode the syllabic rhythm of their native language since early life (Sansavini et al., 1997; Bertoncini et al., 1995) and to rely on such cues to move their first steps into speech segmentation (Nazzi et al., 1998; Nazzi and Ramus, 2003, see Chapter 3).

Specifically, infants seem to rely on the prosodic cues through which relative prominence at the word level is realized in their languages (Jusczyk et al., 1993, 1999). Regarding the present findings, subjects participating in this pilot were Italian-learning speakers; therefore, they have been exposed to a syllable-timed language where patterns of relative prominence are commonly realized by changes in stress at the syllabic level (Nespor et al., 2011). Moreover, word stress in Italian is commonly placed at word-onset, with most of the disyllabic words exhibiting a trochaic stress pattern (d'Imperio and Rosenthall, 1999; White et al., 2009) and polysyllabic words (mostly trisyllabic) showing an ante-penultimate stress (e.g., BI.bi.ta, "drink") or a penultimate stress (Burani et al., 2014, e.g., par.TI.ta, "match"). Previous studies have shown that Italian-learning infants are particularly sensitive to these stress patterns constituting the rhythmic structure of their language and rely on these cues to allocate attentional resources (Russo et al., 2021; Calignano et al., 2021a). Therefore, the increased attentive response to trisyllabic stimuli in this pilot is not likely explainable with a general difficulty in processing polysyllabic linguistic patterns, since they are common in the language background of participants.

An alternative explanation might always take into account the linguistic nature of the stimuli and the structure of the experimental design in considering the general cognitive bias for processing binary over ternary meters in Western infants. Specifically, a binary (musical) meter might be more effective in eliciting the priming effect on matching linguistic stimuli than a ternary one. Accordingly, although the ternary prime block of this Study has not yet been tested, previous findings on the rhythmic priming paradigm in the auditory sensory modality go in this direction (Cason et al., 2015a). Specifically, Cason and Schön (2012) showed that a binary but not a ternary meter was effective in eliciting a priming effect on a behavioral linguistic task. However, the authors also report neural evidence (i.e., a significantly more negative ERP component in the 100–250 ms latency window) for mismatching linguistic stimuli following both the binary and ternary prime, concluding that speech processing was enhanced by the temporal expectations generated by the prime at the neural but not at the behavioral level (Cason and Schön, 2012).

Therefore, proceeding with the counterbalanced design planned in the present Study is essential to contribute to clarifying the effect of temporal regularity on language processing and to extend previous findings on the auditory domain with original evidence on VTS sensory modalities. In fact, the preliminary results from this pilot strongly encourage further research on the effect of VTS rhythms on language processing in infancy. Moreover, the present pilot represents to our knowledge the first application of a vibrotactile device for music perception in studying the effects of rhythmic priming on linguistic sequence processing.

Previous studies have shown that a matching, rhythmic prime facilitates language processing at the phonemic (Cason and Schön, 2012; Cason et al., 2015b), syllabic (Fiveash et al., 2020; Gould et al., 2017, 2018), and syntactic level (Chern et al., 2018; Fiveash et al., 2020). Furthermore, the rhythmic priming effect has been seen to occur also across different sensory modalities with cross-modal transfers from auditory to visual (Fotidzis et al., 2018) and visual to auditory sensory modalities (Liu et al., 2012; Gould et al., 2018). However, no studies have previously investigated other sensory modalities. Therefore, even if the reported preliminary results need to be corroborated, the present pilot represents the first step in favour of a rhythmic priming effect from VTS to auditory sensory modalities. Furthermore, it indicates that the nature of the link between rhythmic and linguistic abilities in infant development is not merely correlational but rather functional, thus representing a possible way to train one domain to improve the other. Research on the strength of this relationship is needed to design possible training targeting the underlying mechanisms of language development. Indeed, a possible interpretation of the effects of VTS rhythmic priming on language processing relies on the perceptual and cognitive mechanisms supporting rhythmic processing in musical and linguistic signals (Fiveash et al., 2021; Ladányi et al., 2020).

Specifically, rhythm is known to elicit entrainment processes on endogenous neural oscillations affecting attentional deployment (Hickok et al., 2011; McAuley, 2010, see Chapter 2 and 3). Regarding the priming effect in particular, the Dynamic Attending Theory suggested that neural oscillations are not only known to entrain in phase with rhythmic stimuli but also to experience entrainment effects even after the external stimulus has stopped (Barnes and Jones, 2000; Jones et al., 2002; Kösem et al., 2018; McAuley, 2010, DAT). This endogenous entrainment therefore drives attentional resource allocation by providing expectations on when an upcoming signal is likely to occur in time (McAuley, 2010). The hierarchical structure of rhythm is known to elicit entrainment mechanisms aiding the cognitive processing of different signals including music and speech (see Chapter 2).

The present Study replicates and extends these findings, providing preliminary evidence on VTS rhythmic effect in eliciting metric expectations across modalities. In line with the DAT, the pattern of observed results might indicate a facilitating effect in processing linguistic stimuli elicited by the VTS rhythmic priming. Furthermore, the general nature of these basic rhythmic abilities is further suggested by the cross-modal transfer of rhythmic priming effect from VTS to auditory sensory modalities. Indeed, cross-sensory transfers between these two modalities might be explained by the overlapping nature of auditory and tactile signals based on vibrations, also observed in the previous studies on VTS rhythms in infancy (Ammirante et al., 2013, 2016; Branje et al., 2014, see Study 4.1). Moreover, these preliminary results might be explained considering the core involvement of sensorimotor processing in processing rhythm (Nayak et al., 2022; Stephan et al., 2016). Infants in particular have been found to rely on sensorimotor information to organize rhythmic stimuli (Phillips-Silver and Trainor, 2005), with audio-motor training enhancing rhythmic effects on language processing (Cason et al., 2015a).

In conclusion, the present pilot advances our knowledge on the underlying cognitive mechanisms explaining the rhythmic priming effect on linguistic processing. Indeed, by means of pupillometry measures it was possible to show that processing of mismatching is aided by attentive resources allocation, as suggested by the increase in cognitive effort in response to mismatching stimuli. This might be interpreted as direct evidence of the effect of neural entrainment on attentional deployment and thus cognitive processing, in line with nowadays theoretical frameworks (Fiveash et al., 2021, the PRISM framework) highlighting the core role of neural entrainment and sensorimotor coupling in rhythm processing, crucial to language development.

4.3.5 Conclusions to Study three

The pilot Study here presented was designed to extend the promising findings from Study one and two on VTS rhythmic processing and language development in infancy. Specifically, the nowadays research lines investigating the link between rhythm and language call for experimental evidence on the strength of this relationship and on the impacts it might have on a developmental perspective. Even though the preliminary results presented here need to be extended, the present pilot aimed at answering this call. In fact, by means of the technological advantages of the Infant vibrotactile system, it was possible to test the application of rhythm on an under-explored sensory modality. Moreover, the pupillometry technique allowed for extending the current understanding of the cognitive mechanisms underlying the effects of rhythm on language processing. Lastly, visualizing and modeling data as continuous measures supported the possibility of exploring preliminary results from a sample of just ten subjects. In conclusion, this pilot was meant to be a first, concrete step to support further research on the potential of applying VTS rhythms to promote language development, extending the promising results presented in Studies one and two. Therefore, the steps moved from the investigation of the VTS rhythmic abilities in infancy, their link to language acquisition, and lastly the potential applications for language development will be summarized and discussed in the last section of this Chapter, reporting the general discussion and conclusions to this research on VTS rhythm and language in development.

4.4 Discussion and conclusion to Chapter 4

Based on previous literature supporting the link between musical rhythm and language, the present Chapter aimed at exploring the extent of this relationship by testing vestibular-tactile-somatosensory (VTS) rhythm perception and its link with early linguistic abilities through infancy. Specifically, three studies investigated: i) the discrimination abilities for the underlying meter of VTS rhythms (Study 4.1); ii) the link between VTS rhythmic abilities with prosodic and phonological auditory skills (Study 4.2); and iii) the possibility to aid auditory linguistic processing with a VTS rhythmic priming (Study 4.3). Rhythmic and linguistic abilities were investigated by means of advanced technologies including: i) the Infant vibrotactile system, specifically designed for this project and based on the music technology field; ii) the eye-tracking system, allowing for collecting complementary, behavioral (i.e., looking times) and physiological (i.e., pupillometry) measures of cognitive processing in preverbal infants; and iii) gaze-triggered paradigms for infant research combined with the Infant vibrotactile system and the eye-tracker thus supporting efficient experimental procedures for investigating the developmental abilities of infants across sensory modalities. Lastly, generalized (linear and additive) mixed-models allowed for modeling the effects of continuous variables such as changes in pupil diameter across time, age, and performance indexes across tasks. This multilevel approach supported the experimental investigation presented in this Chapter providing interesting results. For instance, infants have been found to be able to extract the underlying meter of different musical rhythms only presented in the VTS sensory modality. Accordingly, they are able to detect a change in rhythms with a different meter. Moreover, while young infants direct more attentional resources toward familiar VTS rhythmic patterns (i.e., detected as an increase in pupil diameter toward familiar stimuli), older infants dedicate more attentive resources (i.e., increase in pupil diameter) to novel rhythmic meters (see Section 4.1.3).

Interestingly, infants seem to process double over triple meters differently overall. Specifically, double and triple meters are equally processed by young infants (at 7 months of age) but, compared to ternary meters, rhythms with a binary meter seem to elicit a greater response to deviant stimuli with age. At the end of the investigated age window (i.e., 35 months), triple meters elicit the most reduced response to deviant stimuli (see section 4.1.3). Therefore, this result has been interpreted as possible evidence of an enculturation effect for processing familiar metrical structures in Western developing infants. If similar enculturation processes have been found in music development (Hannon and Johnson, 2005; Hannon and Trainor, 2007) as well as in language and face processing (Maurer and Werker, 2014), Study 1 (section 4.1) is the first to our knowledge to bring evidence about enculturation processes for musical meters perceived under the VTS sensory modality in infancy. Cross-cultural studies might shed light on the role of musical background in shaping the rhythmic abilities of infants, by disentangling whether the processing bias for double meters is elicited by culture or by a general bias for binary structures in humans (Ullal-Gupta et al., 2013; Teie, 2016).

In Study 2 (section 4.2), event-related variations in pupil diameter to mispronounced object-label pairs showed the detailed prosodic and phonological features embedded in the lexical representations of infants from 7 to 35 months of age. Specifically, infants were familiarized with three novel object-label pairs then mispronounced for phonological or prosodic features. Pupil diameter increased in extract time window following the presentation of novel prosodic (i.e., stress) and, especially, novel phonological stimuli. Furthermore, attentive resources (Einhäuser, 2017; Beatty et al., 2000; Mathôt, 2018, detected as an increase in pupil diameter) were increasingly directed toward novel stimuli with age, especially for novel phonological features. Accordingly, older participants (35 months) mainly responded to changes in the phonological dimension while young infants (7 months) directed more attention toward changes occurring in the prosodic dimension (see section 4.2.3).

This result is in line with the literature on infant language development showing the core role of prosodic cues in bootstrapping the first steps of language acquisition in young infants; while fine phonological abilities are known to develop with age, with perceptual narrowing mechanisms aiding the processing of native phonological features as a function of experience (Kuhl, 2004; Gervain, 2018; Jusczyk et al., 1999). Moreover, this result sheds light on the underlying cognitive processes supporting early lexical representations by showing that a different load of cognitive effort is required to process different linguistic features based on the developmental stage, thus providing new evidence about the advantages to apply pupillometry as well as continuous variables (i.e., time) through generalized additive mixed models (Baayen et al., 2017; Tamási et al., 2017, 2019).

However, the main research question beyond this Chapter was to investigate the link between VTS rhythmic abilities and linguistic skills in infancy. To this end, data from Study 1 were added as a continuous predictor to a further model accounting for the infant linguistic abilities collected in Study two. Specifically, a rhythmic index was computed by the proportion of pupil dilation toward novel stimuli in the test phase. Model comparison suggested that the best model was the one accounting for the interaction between rhythm and language, together with age. Specifically, infants allocating more attentive resources to familiar rhythms were more likely to direct increasing cognitive resources toward familiar linguistic labels. By contrast, infants allocating more attentive resources toward novel stimuli in the rhythmic task were the same showing a greater attentive response to novel linguistic features. Therefore, general VTS rhythmic abilities seem to be related to early phonological and prosodic skills through infancy. However, the direction of this relationship seems to change and, since the model estimated a significant, additive effect of age, a further model was computed to explore the changes in the relationship between the two domains as a function of age.

This last analysis showed that the amount of attentive resources shifted from familiar toward novel stimuli as a function of age and in a parallel fashion for rhythmic and linguistic stimuli. This result is explained here by considering that learning occurs where attention goes. That is, allocating attentive resources toward a given stimulus maximizes the processing thus supporting learning.According to this view, a robust representation of familiar stimuli is required to produce a novelty response. That is, infants who have fully processed familiar stimuli are ready to allocate newly available attentive resources to the processing of new information. This principle has been previously described in the Hunter and Ames (1988) developmental model of attention. Specifically, it is hypothesized that infant attention to a stimulus is dictated by the stability of the representation in memory. This can occur as a function of stimulus complexity (i.e., simple stimuli require less familiarization times to elicit a novelty response; e.g., Cornell, 1975; Martin, 1975), familiarization time (Martin, 1975; Colombo and Bundy, 1983, longer familiarization time might more likely elicit a novelty response), or age (Roder et al., 2000, familiarity responses are more commonly observed early in development).

Therefore, results in Study one and two might be explained considering familiarity and novelty responses as reflecting different levels of expertise across age. Specifically considering the nature of the stimuli here investigated, expertise in phonological processing is observed to grow with age with a peak compatible with the naming explosion (around 18 months, see Figure 4.13; Barrett, 2017). Furthermore, perceptual tuning toward predominant stimuli in the infant environment is known to occur at around after the end of the first year in many domains including rhythm perception (Hannon and Trehub, 2005). Specifically, infants at 6 months were found to display culture-general responses whereas an adult-like, culture-specific pattern of responding to musical rhythms seems to emerge with age.

Therefore, the findings presented in Study one and two represent an extension of previous research on perceptual and attentive mechanisms underlying rhythm and speech processing in infancy. Specifically, the main advance of this work is to include the VTS sensory channel to more largely investigated visual and auditory modalities. By doing so, it has been observed that infants show discriminative skills for VTS rhythms which might follow the perceptual narrowing dynamics of other sensory domains. Moreover, it has be shown that the infant ability to successfully extract and encode the underlying structure (i.e., meter) of a VTS rhythmic stimulus is related to the ability to successfully detect linguistic changes occurring on the prosodic and phonological structures of labels. Lastly, a developmental change has been observed in the response patterns of both the single studies and on their relationship. Specifically, a familiarity response emerged in younger participants who allocate more attentive resources toward familiar, rhythmic and linguistic, patterns; while older infants displayed the opposite trend with an increased resource allocation toward novel, rhythmic and linguistic, structures. Therefore, it might be concluded that results are in line with the hypothesis of a link between VTS rhythmic and linguistic abilities in infancy.

Inspired by this promising result, Study three (section 4.3) was designed to test a possible application of the Infant vibrotactile system in investigating the direct effects of VTS rhythms on language processing. A pilot for this Study was run and presented in this Chapter. The design of this experiment was based on rhythmic priming paradigms but including the novel research question of whether cross-modal effects (i.e., VTS to auditory) were possible from one domain (i.e., music) to another (i.e., language). Preliminary results on ten participants are encouraging, displaying a facilitating effect of the rhythmic prime on processing matching, linguistic stimuli. As in the two previous studies, event-related variations in pupil diameter over time were taken as an index of cognitive load, contributing in advancing the experimental paradigm by means of a novel measure, as well as the state of the art on the underlying cognitive processing supporting the rhythmic priming effect. Specifically, the role of attentive resources allocation was discussed within the Dynamic Attending Theory, thus suggesting that the entrainment of neural oscillations to an external rhythm drives the attentive system toward the incoming stimulus structure by providing expectation on when an upcoming signal is likely to occur in time (McAuley, 2010; Jones, 2018).

Therefore, even if these preliminary results need further extension, Study three moved a first step into possible future applications of the Infant vibrotactile system in bootstrapping language development. A further application of this system might regard sensory-deprived populations, as early deaf or/and blind infants. Deaf infants in particular might benefit from a vibrotactile exposure since a fully recognized (Pisoni et al., 2016, still poorly understood) difficulty for these children concerns domain-general sequencing abilities (Conway et al., 2009). Specifically, a lack of experience with sound early in life might produce a poor ability to successfully organize events in time thus impacting the processing of several inputs (including but not limited to language). Accordingly, a significant variance in the outcomes following cochlear implantation is observed especially during school years when read-writing abilities are acquired (Pisoni et al., 2016). Since rhythm is a powerful tool for the temporal encoding of sequential signals, providing this experience early in life through the VTS sensory channel could help in bootstrapping general abilities to sustain cognitive development. A first attempt to test this possible application has started, in collaboration with the Audiology Center in the University Hospital of Padua.

4.4.1 Conclusions

In conclusion, this Chapter provided original evidence on the link between the VTS rhythmic experience and language acquisition in infancy. This contribution might be useful in informing the nowadays models of cognitive development (Fiveash et al., 2021; Ladányi et al., 2020, the PRISM framework; the Atypical rhythm risk hypothesis). Furthermore, the application of technological advances in music technology (i.e., the Infant vibrotactile system) might enrich the infant cognitive as well as the music technology research field. Third, the use of a multilevel approach in this Study might motivate researchers from the field in applying pupillometry to classical looking times to better focus on the investigated processes (Sirois and Jackson, 2007; Jackson and Sirois, 2009). Furthermore, modeling age as a continuous predictor in mixed-model rather than dividing the sample in classical, categorical age windows allowed for the emergence of interesting developmental trajectories. Therefore, this corpus of work might be useful in contributing to improve research methodologies and practices in infant research. The latest impact regards the application of the present findings in clinical and educational contexts. Specifically, the strength of the relationship between general rhythmic skills and linguistic development need to be explored to inform clinicians and health professionals about possible risk factors related to rhythm processing.

Therefore, the present research aimed at representing a small step toward this direction. By replicating and extending these findings, maybe one day early screening will include the rhythmic abilities while training practices including rhythm and multi-sensory perception will be applied, not only in the case of at-risk or atypical populations but also to enrich educational practices. Accordingly, school ages represent a critical challenge for many children who struggle in acquiring how to read and write or to concentrate. Even though the effect of the prenatal experience with VTS rhythms is faraway at that point, the embodied perception of rhythm is still in place and influences perception and cognition (Phillips-Silver and Trainor, 2007). Therefore, active sensorimotor experiences might serve as a potential channel to generate positive effects on higher cognitive abilities. For this reason, the effect of a sensorimotor rhythmic stimulation on reading and writing abilities will be evaluated in young adults in the last two Studies of this thesis, as described in the next Chapter.

Chapter 5

Sensorimotor rhythm in literacy

In the previous Chapters, the role of VTS rhythmic experiences in early language development has been explored and discussed. However, the development of language is a long-life process in which individuals grow and mature their skills by building upon the basic abilities they have acquired in previous steps. Thus, they become more and more proficient in the use of language. Following this developmental path, language acquisition grows in complexity from basic auditory skills, foundational to prosodic and phonological processing in infancy, to higher cognitive capacities supporting the acquisition of literacy abilities in school ages. In the present Chapter, the contribution that the current research on rhythm processing might provide to more advanced stages of language development will be explored. Specifically, the difficulties experienced by individuals with dyslexia will be targeted, based on the growing body of research suggesting a possible role of general rhythmic skills in explaining the linguistic deficits in this population (Flaugnacco et al., 2015; Forgeard et al., 2008; Goswami, 2002; Goswami et al., 2013; Huss et al., 2011; Lee et al., 2015; Thomson and Goswami, 2008; Dellatolas et al., 2009; Muneaux et al., 2004; Goswami, 2011; Ladányi et al., 2020; Lense and Dykens, 2016; Nayak et al., 2022). Specifically, the TSF by Goswami (2011) proposed that an atypical synchronous activity in the oscillating networks at different frequency bands (Buzsaki, 2006, Delta, 1.5–4 Hz; Theta, 4–10 Hz; and Gamma, 30–80 Hz) might explain the auditory difficulties which lead to phonological impairments in dyslexia (Goswami, 2011). Consistently, several studies have reported atypical neural entrainment and rhythm processing in dyslexia (Leong and Goswami, 2014; Molinaro et al., 2016; Goswami et al., 2010; Surányi et al., 2009), leading to the hypothesis that atypical rhythm might be targeted as a risk factor of language impairment (Ladányi et al., 2020; Lense et al., 2021; Nayak et al., 2022, see Chapter 3).

Interestingly, rhythmic difficulties have been proposed to occur across sensory modalities, since linked to a general impairment in neural phase locking leading to cascade effects on cognitive and behavioral functions (Goswami, 2011). Consistently, general sluggish attention-shifting theories report impairments in attentive cueing paradigms when processing both visual and auditory stimuli (Hari and Renvall, 2001; Facoetti et al., 2010). However, less is known about other sensory modalities and, especially, about cross-sensory dynamics. In fact, beyond visual and auditory processing, what is currently known mainly refers to the motor domain only. Specifically, individuals with dyslexia have been found to show difficulties in finger-tapping tasks (Wolff, 2002; Colling et al., 2017; Persici et al., 2019; Thomson et al., 2006; Thomson and Goswami, 2008). Since a specific motor deficit in dyslexia has not been proved (Colling et al., 2017; Kronbichler et al., 2002; Ramus et al., 2003), motor difficulties in this population are hypothesized to be caused by a general timing impairment, paralleling the findings from the auditory and visual functioning (Wolff, 2002; Goswami, 2011).

Therefore, a sensorimotor lent that investigates performances across both sensory (e.g., auditory, visual, or tactile) and motor domains has been proposed as a useful tool for understanding the timing deficits in dyslexia (Ramus et al., 2003). Furthermore, interventions focusing on rhythmic training have been proved to be effective in supporting language development in dyslexia with similar evidence resulting from short-term exposures to rhythm in priming tasks (Bonacina et al., 2020; Flaugnacco et al., 2015; Overy, 2003; Frey et al., 2019; Bouloukou et al., 2021; Cancer et al., 2019; Ostiz-Blanco et al., 2021; Germagnoli et al., 2016; Ladányi et al., 2021; Przybylski et al., 2013; Bedoin et al., 2016; Canette et al., 2019, 2020). Nevertheless, the underlying mechanisms sustaining the effect of a rhythmic training on language processing in dyslexia need to be fully understood, especially, the effect of long-term and short-term exposure to rhythm need to be further explored. In fact, if the literature agrees on the effectiveness of long-term training, the improvements resulting from direct, short-term exposure to rhythm are still an open debate. That is if individuals with dyslexia struggle in synchronizing to an external beat, what are the mechanisms through which a regular rhythmic cue enhances their performances in linguistic tasks?

Furthermore, how to maximize the focus on sensorimotor mechanisms thus studying timing processing across sensory modalities? And lastly, at which level of the cascade of sensory, cognitive, and behavioral events supporting the efficiency of literacy skills might rhythm more likely have an impact? To answer these questions, the present Chapter illustrates two studies investigating the literacy abilities of young adults. Specifically, the effects of sensorimotor rhythm on phonological processing will be evaluated in the reading (Study 4) and writing (Study 5) abilities of 60 students. Moreover, self-report literacy difficulties, years of music education, and working memory abilities will be investigated. Results will be illustrated and discussed in terms of the possible impacts of sensorimotor rhythmic interventions on the underlying sensory and processing mechanisms supporting language abilities across development.

5.1 Study 4: influence of sensorimotor rhythm on reading abilities

In this Study, the effect of regular rhythmic stimulation on reading will be evaluated in sixty, healthy young adults. Specifically, participants will be asked to perform a lexical decision task for word and non-word in three different conditions: i) under a baseline setting, ii) under articulatory suppression, and iii) under articulatory suppression with a regular rhythmic cue. Specifically, the lexical decision task (Coltheart et al., 1979), as performed in the baseline condition, is a classical experimental paradigm in which participants are requested to classify visually presented, word and non-word items. To perform this task, verbal information must be encoded and stored in the phonological loop which requires working memory and articulatory rehearsal (Baddeley et al., 1984). Commonly, working memory is measured by means of digit span for verbal items (Reisberg et al., 1984, digits, letters, and words), whereas articulatory rehearsal can be assessed under articulatory suppression (Li et al., 2021). Articulatory suppression is the interfering effect produced on the phonological loop by the articulation of irrelevant information during a verbal task (Baddeley, 1993). In the present Study, articulatory suppression is applied to increase the cognitive demand required in performing the lexical decision task, thus letting possible difficulties in the phonological processing underlying reading abilities emerge. Indeed, current screening for dyslexia assessing reading and writing abilities in young adults commonly apply articulatory suppression (Martino et al., 2011; Sartori et al., 1995; Colombo et al., 2009). Crucially, in the present Study a third condition was added to the classical reading assessments by presenting participants with an auditory, regular rhythmic cue. Specifically, the rhythmic cue consisted of a metronome signal played at 2 Hz. Although syllabic rate across languages is approximately 200 ms, the 2 Hz frequency is thought to be crucial in language processing since it reflects the rhythm of stressed syllables (Arvaniti, 2009). Moreover, a low frequency of 2 Hz elicits a Delta band process in the brain which was found to be impaired in individuals with dyslexia (Thomson et al., 2006; Thomson and Goswami, 2008).

Therefore, according to the TSF which posits a timing deficit underlying phonological impairments in dyslexia (Goswami, 2011), presenting participants with a regular rhythmic cue at 2 Hz should facilitate their performances in the reading task, containing the interfering effects of the articulatory suppression. Moreover, the effect of sensorimotor abilities will be evaluated since participants will be presented with visual (i.e., linguistic items) and auditory (i.e., rhythmic cue) sensory stimulation whereas they will be asked to vocally articulate sounds for suppression. Performances in the three conditions will be evaluated in terms of response accuracy and cognitive effort, measured through pupillometry (Beatty et al., 2000; Mathôt, 2018; Einhäuser, 2017). Moreover, the investigation of the phonological loop will be enhanced by adding a digit span task to assess working memory. The evaluation of the effect of rhythm will be completed by investigating and quantifying any previous long-term exposure to rhythm in years of music education. Lastly, participants will be asked to report any experienced difficulty in reading and writing through a standardized self-report questionnaire (Vinegard, 1994).

5.1.1 Aim and research questions

The aim of this study is to extend previous knowledge on the influence of rhythm on the phonological mechanisms underlying reading abilities in young adults. Specifically, the sensorimotor effect of entrainment to an external rhythm will be evaluated by presenting participants with a regular rhythmic cue while performing a lexical task under articulatory suppression. According to the TSF (Goswami, 2011), a regular rhythmic cue might facilitate the linguistic processing of stimuli thus containing the disrupting effect of articulatory suppression on phonological processing. This would occur if the neural oscillations do entrain to the regular cue thus aiding the allocation of attentional resources in time. If so, then the phonological loop would be less affected by the interfering impact of the articulatory suppression. The role of working memory could then be directly linked to a more efficient temporal allocation of attention leading to cascading improvements in phonological processing. If so, then higher performance in working memory assessed through the digit span task (i.e., scores in the digit task) should predict better performance in the linguistic task.

Furthermore, the cross-domain effect of rhythm on language should be further confirmed by a positive association between any previous experience in music education and linguistic performances. Therefore, if rhythm does provide a facilitating effect in the underlying mechanisms sustaining phonological processing in reading, then an improvement in performance should be expected for those individuals who struggle the most with reading skills. By contrast, the timing deficit in individuals with reading difficulties could be so substantial to prevent the possibility to benefit from sensorimotor entrainment, thus lowering performances. Therefore, results from this Study might contribute to: i) testing whether a regular rhythmic cue can sustain the phonological processes underlying reading abilities in young adults, ii) whether receiving a music education predicts better reading performances, and iii) whether individuals with reading difficulties benefit from a regular rhythmic cue or whether they show impaired rhythmic processing.

5.1.2 Method

Participants

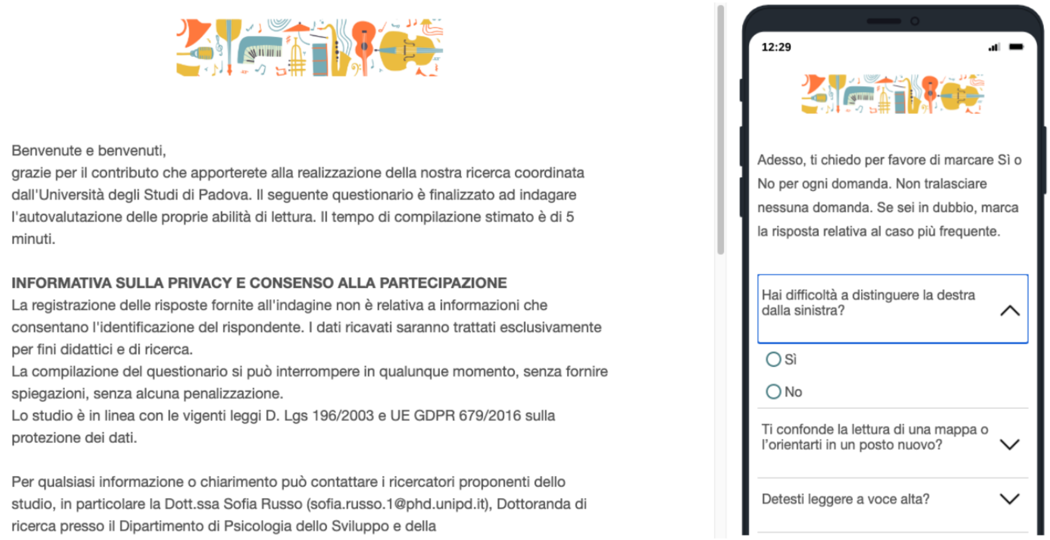
Sixty, healthy young adults (37 females, M = 21 age, SD = 2) participated in this Study. All participants were students of the University of Padua and were invited to participate and visit the eye-tracking laboratory of the Department of Developmental Psychology and Socialization. The inclusion criteria for all participants were: to be in good health, to have no sensory or neurological disorders or any familiar language disorder, and to be native Italian speakers. None of the participants were professional musicians. The research protocol was approved by the Ethics Committee of the Department of Developmental Psychology and Socialization, University of Padua (protocol number: 2423). The research was conducted in accordance with the principles of the Declaration of Helsinki. Participants signed a consent form before starting the experimental section. In the consent form, they agreed to participate in a second experimental section too, as part of an extended project on literacy efficiency in young adults (see Study 5.2).

Stimuli and Materials

Verbal stimuli for the lexical decision task consisted of 48 items from the Italian Battery for the assessment of dyslexia and developmental dysorthography by Sartori et al. (1995). Original stimuli were divided into three lists of 16 items each, balanced for word (n=8) and non-word (n=8). Moreover, the number of phonemes in each item was balanced across lists per word and non-word. Stimuli for the digit span test consisted of 16 lists of digits, auditory presented from the Digit Span Forward subtest of the Wechsler Adult Intelligence Scale, Fourth Edition (Hartman, 2009, WAIS-IV). Rhythmic stimulation was provided via a metronome paced at 2 Hz (120 bpm). To balance the presence of auditory stimulation across conditions, white noise was played in the background. Difficulties in reading and writing were investigated by means of a self-report questionnaire consisting of 20 items with yes/no answers (Vinegard, 1994, A revised Dyslexia Checklist). Two items asking whether participants received, and for how many years, individual lessons with a musical instrument were added to the checklist to investigate music education (Zhang et al., 2020).

Apparatus

The lexical decision task was re-adapted for a computerized administration, thus programmed, and presented through Open Sesame software 3.1 (Mathôt et al., 2012) on a computer laptop (Acer travel mate 5772 g). Lexical stimuli were visually presented on a 27-inch monitor (Philips 300 × 300). A remote, infrared eye-tracking camera (Tobii X2-60 Eye-Tracker) placed directly below the screen recorded the eye movements using bright-pupil technology at a sampling frequency of 60 Hz. Behavioural responses were collected through a keyboard placed below the monitor, in front of the participant. The dyslexia checklist, with the additional items on music education, was administered online by means of the Qualtrics platform (Qualtrics, Provo, UT; Figure 5.2).



Benvenute e benvenuti,
grazie per il contributo che apporterete alla realizzazione della nostra ricerca coordinata dall'Università degli Studi di Padova. Il seguente questionario è finalizzato ad indagare l'autovalutazione delle proprie abilità di lettura. Il tempo di compilazione stimato è di 5 minuti.

INFORMATIVA SULLA PRIVACY E CONSENSO ALLA PARTECIPAZIONE
La registrazione delle risposte fornite all'indagine non è relativa a informazioni che consentano l'identificazione del rispondente. I dati ricavati saranno trattati esclusivamente per fini didattici e di ricerca.
La compilazione del questionario si può interrompere in qualunque momento, senza fornire spiegazioni, senza alcuna penalizzazione.
Lo studio è in linea con le vigenti leggi D. Lgs 196/2003 e UE GDPR 679/2016 sulla protezione dei dati.

Per qualsiasi informazione o chiarimento può contattare i ricercatori proponenti dello studio, in particolare la Dott.ssa Sofia Russo (sofia.russo.1@phd.unipd.it), Dottoranda di ricerca presso il Dipartimento di Psicologia dello Sviluppo e della

Adesso, ti chiedo per favore di marcare Sì o No per ogni domanda. Non tralasciare nessuna domanda. Se sei in dubbio, marca la risposta relativa al caso più frequente.

Hai difficoltà a distinguere la destra dalla sinistra?
○ Sì
○ No

Ti confonde la lettura di una mappa o l'orientarti in un posto nuovo?

Detesti leggere a voce alta?

FIGURE 5.1: Online administration of the revised dyslexia checklist. Screens of the online rendering for devices (i.e., computer/iPad on the left and smartphone on the right) of the qualtrics.com platform displaying part of the consent form (left) and part of the checklist items (right).

Experimental design and procedure

During the recruitment process, participants were asked to fill in the online questionnaire prior to the experimental section. Once accessed the form via a link or QR code, the first screen stated the consent form and asked for approval, then the checklist appeared. The experimental section was carried out at the eye-tracker laboratory at the Department of Developmental Psychology and Socialization of the University of Padua. After signing the consent form, participants were presented with the digit span test first. Instruction and administration were consistent with the Digit Span Forward subtest of the WAIS-IV test (Hartman, 2009). Therefore, individual scores were obtained from the longest string of digits repeated without error over two trials (Greiffenstein et al., 1994). After a short break, participants were tested with the lexical decision task. The eye-tracker calibration was run at the beginning of the test. Participants were presented with fixation points at five different locations on the screen (i.e., top-left, top right, centre, bottom-left and bottom-right). Once the calibration was successfully completed, the experimental section began with two training trials presented to familiarize participants with the task.

The experimental design consisted of three conditions: baseline, suppression, and rhythm (see Figure 2). In the baseline condition, the lexical decision task was performed with no interference or cue. In the suppression condition, participants were instructed to perform the lexical decision task under articulatory suppression, thus continuously repeating the syllable La while completing the task. In the rhythm condition, participants were cued by a metronome in performing the lexical decision task under articulatory suppression. To balance the presence of auditory stimulation across conditions, white noise was played in the background. In each condition, one of the three lists of items was presented. The experimental design was counterbalanced based on i) order presentation, and ii) list of items. Specifically, half of the participants were presented with the baseline condition first, then the suppression condition, and the rhythmic condition in the end, whereas the other half of participants was presented with the opposite order. Moreover, the lists of items assigned to the suppression and rhythmic conditions were balanced for both orders resulting in 2x2 final combinations, each completed by 15 participants.

Within each list, trials were presented in randomized order. During the experimental section, participants were seated on a chair positioned 60 cm away from the monitor and were tested individually in a quiet room. Semi-darkness constant luminance was guaranteed by the light lining of the eye-tracker laboratory.

Statistical analyses

Performances on the lexical task were coded in terms of accuracy (i.e., the proportion of correct over total responses) and cognitive load (Beatty et al., 2000; Einhäuser, 2017; Mathôt, 2018; Calignano et al., 2021b, i.e., changes in pupil size under constant luminance for correct trials). As mentioned above, working memory scores were obtained from the longest string of digits repeated without error over two trials in the digit span test (Hartman, 2009; Greiffenstein et al., 1994, from 2 to 9). Self-reported difficulties in literacy abilities were coded as positive answers in the checklist for dyslexia (Vinegard, 1994, from 0 for no difficulties to 20 for significant difficulties). Long-term exposure to rhythm was coded as years of individual lessons with a musical instrument (from 1 to 10). Behavioural (i.e., accuracy) and physiological (i.e., variations in pupil diameter) data from the lexical task were modelled as a continuous variable with gamma and gaussian distribution respectively, by means of Generalized Mixed-Effects models (Nieuwenhuis et al., 2012; Fox, 2015). Pre-processing steps (i.e., transforming the raw data into a format for statistical analysis and visualization) were performed following the Mathôt and Vilotijević (2022) guidelines. Since the pre-processing procedure here applied is the same as in Study 4.1, 4.2, and 4.3, please refer to section 4.4 in Chapter three for further specifications. The fixed effects of conditions (i.e., block) and class (i.e., word/non-word) were gradually added to the null model accounting for participant and item as random effect. The resulting best model was then updated to investigate the effect of: i) literacy difficulties, ii) music education, and iii) working memory efficiency. To investigate the predictive value of each effect, a model comparison was performed using the likelihood ratio test (LTR), Akaike Information Criterion (Akaike, 1974, AIC), and AIC weight as indexes of the goodness of fit. The resulting models were inspected one by one (sections from 5.1.3 to 5.1.3). Then, a model comparison was performed to estimate the best predictor between i) literacy difficulties, ii) music education, and iii) working memory efficiency (section 5.1.3).

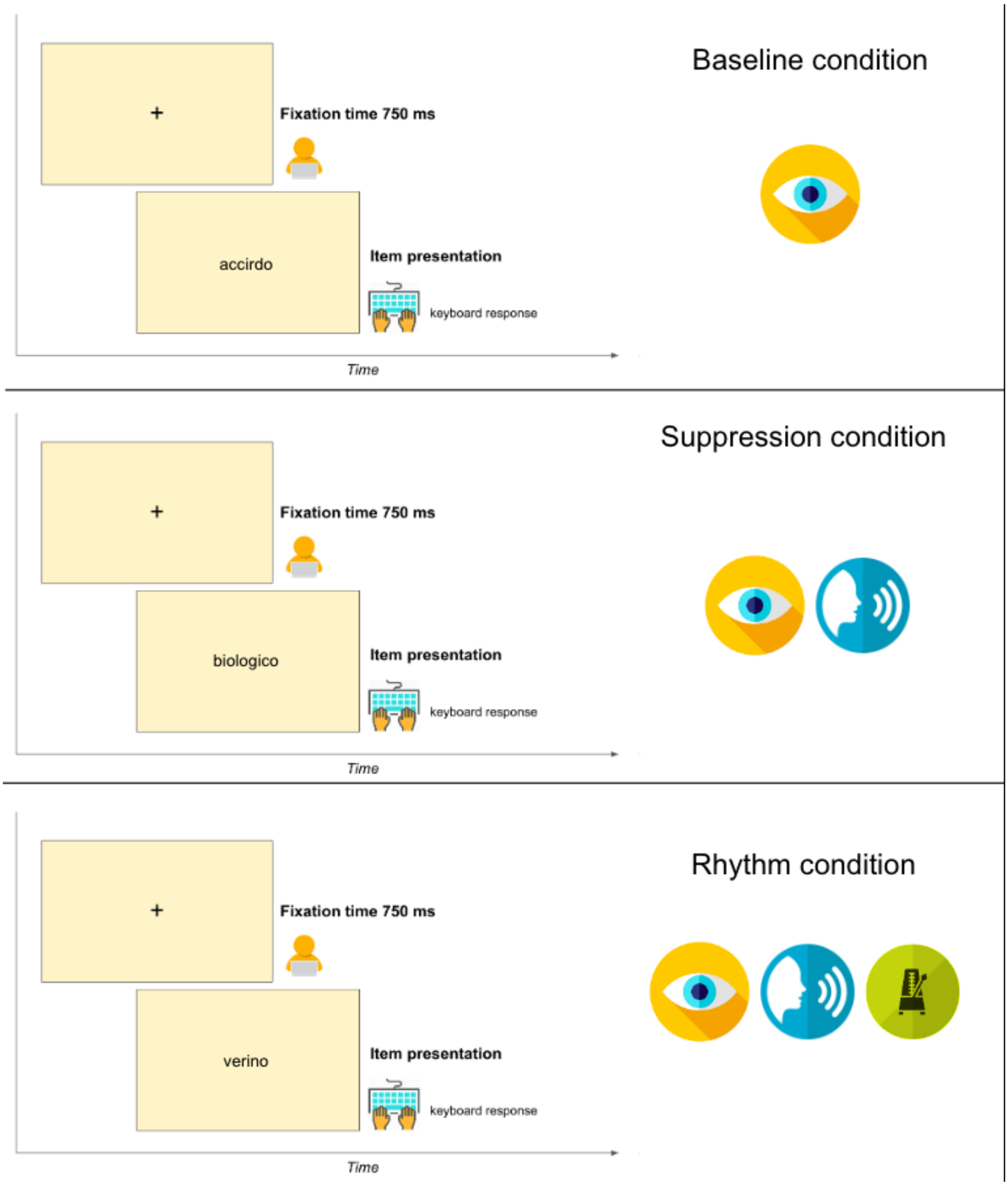


FIGURE 5.2: Experimental design from Study 4. For each condition, participants were instructed to perform the lexical decision task (left). For each trial, a fixation cross appeared on the screen and remained visible for 750 ms. Then, a lexical item was presented, and participants were instructed to answer by keyboard pressing. Therefore, the baseline condition mainly interested the visual sensory modality. In the suppression condition, participants were instructed to perform the task under articulatory suppression, thus vocally producing a verbal output (i.e., *La*) while reading the word and non-word items. In the rhythmic condition, a regular rhythmic cue (i.e., a metronome paced at 2 Hz) assisted participants in performing the lexical decision task under articulatory suppression.

5.1.3 Results

Literacy difficulties

The model comparison showed that M5, accounting for the interaction between the effects of condition, class, and literacy difficulties (coded as scores in the self-reported checklist for dyslexia screening; see section 5.1.2) resulted in the best model in explaining performance during the lexical task in terms of both accuracy (Table 5.1) and cognitive load (Table 5.2). Specifically, Model 5 predicted a significant decrease in accuracy for participants with high self-report difficulties in literacy performing lexical decisions on non-words during the suppression and the rhythmic conditions compared to the baseline (suppression: b=-0.0106, SE=0.0004, t=-22.723, p<0.001; rhythm: b=-0.0146, SE=0.0005, t=-29.215, p<0.001). Regarding pupillometry, the model predicted a significant increase in pupil diameter for participants with high self-report difficulties in literacy performing lexical decisions on non-words during the suppression (b=0.0049, SE=0.0004, t=10.267, p<0.001) but not the rhythmic (b=0.00007, SE=0.0005, t=0.138, p=0.891) condition. Descriptive statistics and predicted effects are shown in Figure 5.3 and 5.4, respectively.

TABLE 5.1: GLMM comparison for literacy and accuracy in Study 4

Models	Deviance	dAIC	AICw
M.0 Accuracy ∼(1 ∣ id) + (1 ∣ item)	1241	11380	0
M.1 Accuracy ∼condition + (1 ∣ id) + (1 ∣ item)	1234	10856	0
M.2 Accuracy ∼condition + class + (1 ∣ id) + (1 ∣ item)	1225	10122	0
M.3 Accuracy ∼condition * class + (1 ∣ id) + (1 ∣ item)	1202	8239	0
M.4 Accuracy ∼condition * class + literacy_difficulties + (1 ∣ id) + (1 ∣ item)	1202	8240	0
M.5 Accuracy ∼condition * class * literacy_difficulties + (1 ∣ id) + (1 ∣ item)	1106	0	1

TABLE 5.2: GLMM comparison for literacy and cognitive load in Study 4

Models	Deviance	dAIC	AICw
M.0 Pupil ∼(1 ∣ id) + (1 ∣ item)	-1656	435	0
M.1 Pupil ∼condition + (1 ∣ id) + (1 ∣ item)	-1657	371	0
M.2 Pupil ∼condition + class + (1 ∣ id) + (1 ∣ item)	-1657	339	0
M.3 Pupil ∼condition * class + (1 ∣ id) + (1 ∣ item)	-1657	354	0
M.4 Pupil ∼condition * class + literacy_difficulties + (1 ∣ id) + (1 ∣ item)	-1657	369	0
M.5 Pupil ∼condition * class * literacy_difficulties + (1 ∣ id) + (1 ∣ item)	-1661	0	1

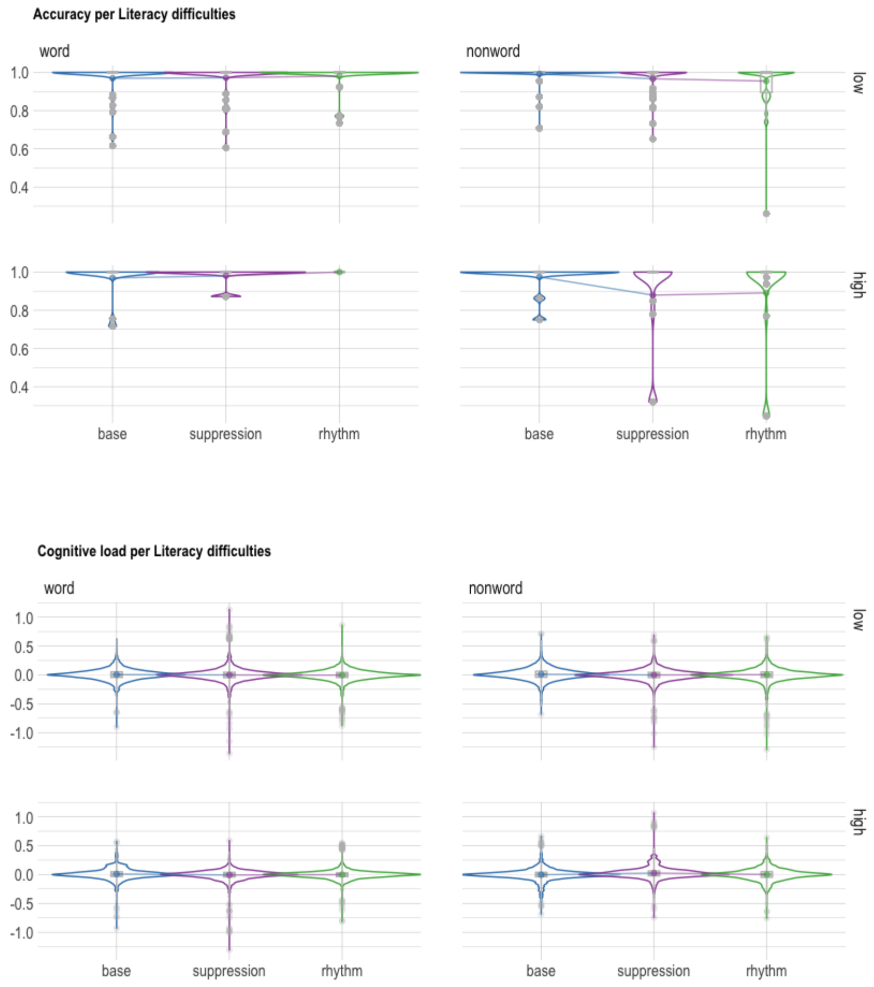


FIGURE 5.3: Descriptive statistics for literacy difficulties. Data are reported for accuracy (top) and baseline corrected pupil (bottom). For this plot, participants were divided into proficient or poor readers (out-range scores in the self-reported checklist for literacy difficulties). Data are shown for conditions (i.e., baseline, suppression, and rhythm) and class (i.e., word, non-word).

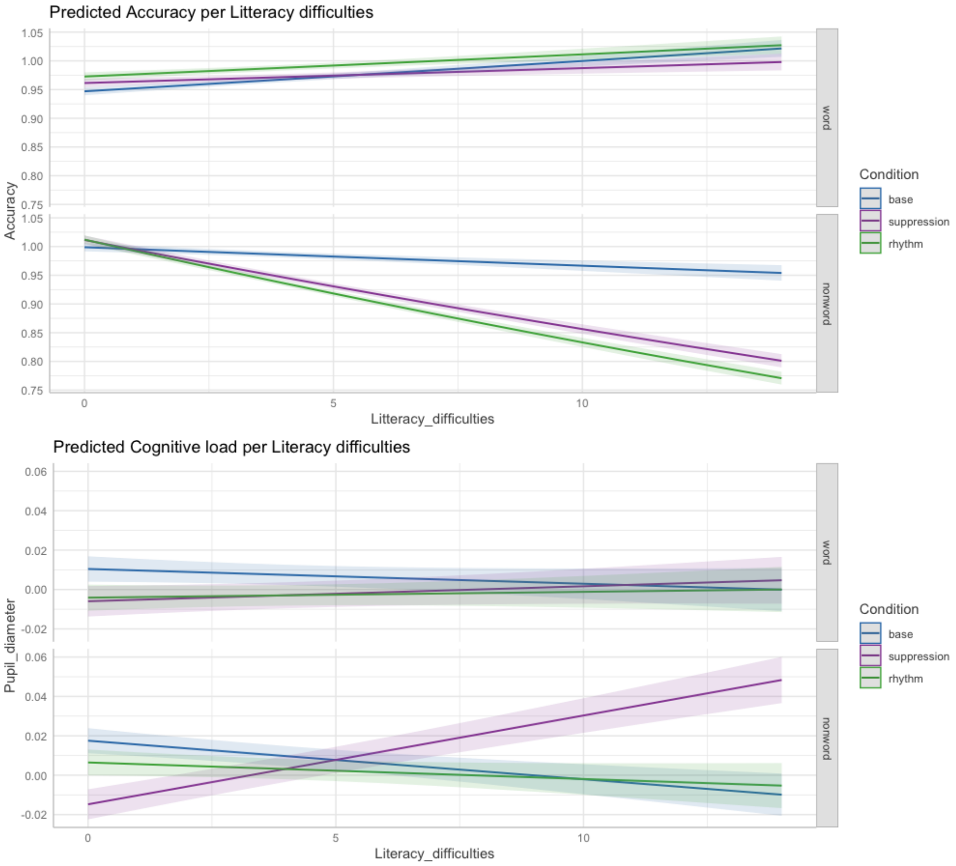


FIGURE 5.4: Effect plots of the best model (M5). The interaction effect of condition (i.e., baseline, suppression, and rhythm), class (i.e., word and non-word), and literacy difficulties (i.e., self-report difficulties in everyday literacy skills) is here plotted on predicted values of Accuracy (i.e., proportion of correct over total answers) and cognitive load (i.e., increase in pupil size under constant luminance).

Music education

Model comparison showed that M5, accounting for the interaction between the effects of condition, class, and years of music experience (see section 5.1.2) resulted as the best model in explaining performance during the lexical task in terms of both accuracy (Table 5.3) and cognitive load (Table 5.4). Specifically, Model 5 predicted a significant increase in accuracy for participants with prior musical education performing lexical decisions on non-words during the suppression and the rhythmic conditions (suppression: b=0.0087, SE=0.0007, t=12.138, *p*<0.001; rhythm: b=0.0230, SE=0.0007, t=31.636, *p*<0.001). Regarding pupillometry, the model predicted a significant increase in pupil diameter for participants with prior musical education performing lexical decisions on non-words during the suppression condition (b=0.0049, SE=0.0004, t=10.267, *p*<0.001) and a decrease in the rhythmic condition (b=-0.0021, SE=0.0007, t=-3.022, *p*=0.003). Overall, music education predicted less cognitive load in the rhythmic condition (b=-0.0020, SE=0.0005, t=-3.736, *p*<0.001). Descriptive statistics and predicted effects are shown in Figure 5.5 and 5.4, respectively.

TABLE 5.3: GLMM comparison for music education and accuracy in Study 4

Models	Deviance	dAIC	AICw
M.0 Accuracy ∼(1 | id) + (1 | item)	1241	4924	0
M.1 Accuracy ∼condition + (1 | id) + (1 | item)	1234	4399	0
M.2 Accuracy ∼condition + class + (1 | id) + (1 | item)	1225	3665	0
M.3 Accuracy ∼condition * class + (1 | id) + (1 | item)	1202	1782	0
M.4 Accuracy ∼condition * class + music_education + (1 | id) + (1 | item)	1202	1784	0
M.5 Accuracy ∼condition * class * music_education + (1 | id) + (1 | item)	1180	0	1

TABLE 5.4: GLMM comparison for music education and cognitive load in Study 4

Models	Deviance	dAIC	AICw
M.0 Pupil ∼(1 | id) + (1 | item)	-1656	115	0
M.1 Pupil ∼condition + (1 | id) + (1 | item)	-1657	51	0
M.2 Pupil ∼condition + class + (1 | id) + (1 | item)	-1657	19	0
M.3 Pupil ∼condition * class + (1 | id) + (1 | item)	-1657	34	0
M.4 Pupil ∼condition * class + music_education + (1 | id) + (1 | item)	-1657	48	0
M.5 Pupil ∼condition * class * music_education + (1 | id) + (1 | item)	-1657	0	1

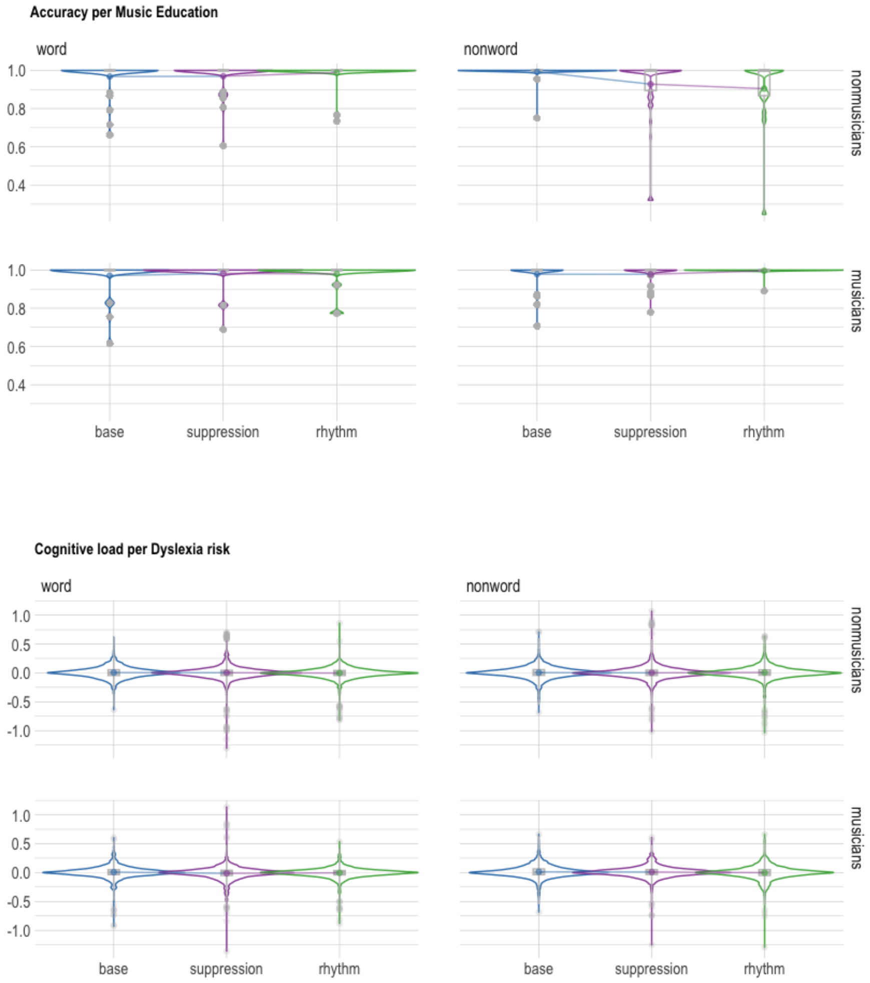


FIGURE 5.5: Descriptive statistics for music education. Data are reported for accuracy (top) and baseline corrected pupil (bottom). For this plot, participants were divided into musicians and non-musicians (based on whether they received music training or not). Data are shown for conditions (i.e., baseline, suppression, and rhythm) and class (i.e., word, non-word).

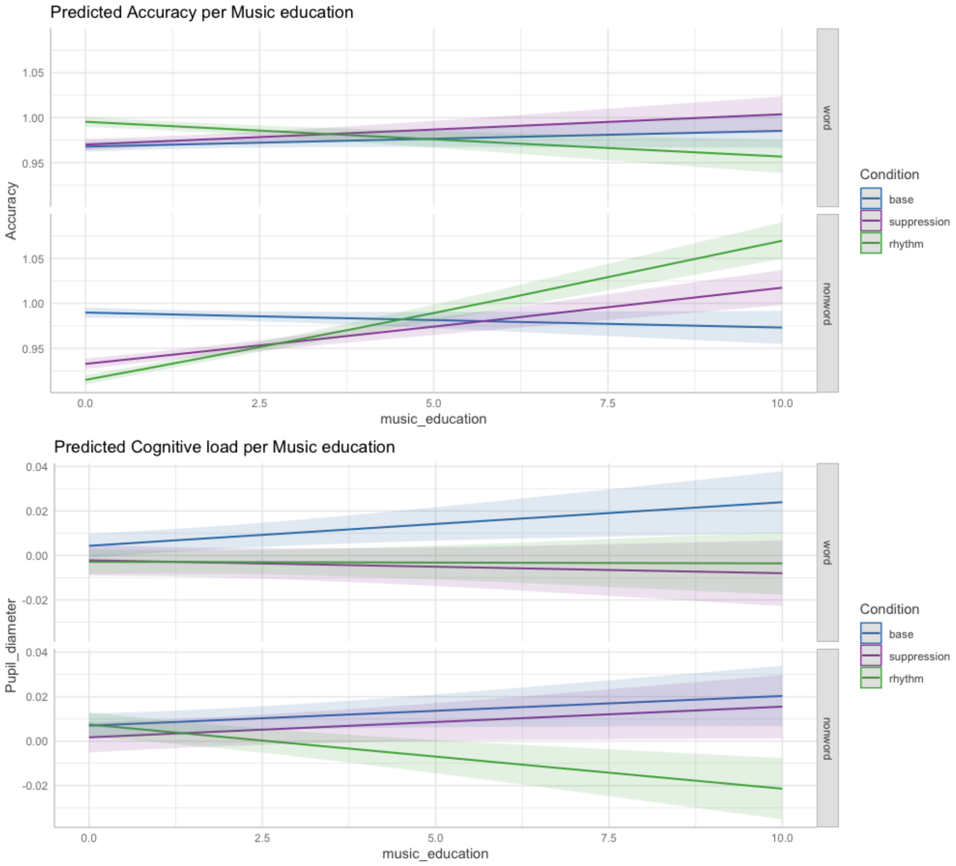


FIGURE 5.6: Effect plots of the best model (M5). The interaction effect of condition (i.e., baseline, suppression, and rhythm), class (i.e., word and non-word), and music education (i.e., number of years) is here plotted on predicted values of Accuracy (i.e., proportion of correct over total answers) and cognitive load (i.e., increase in pupil size under constant luminance).

Working memory

The model comparison showed that M5, accounting for the interaction between the effects of condition, class, and scores in the digit span task (see section 5.2.2) resulted as the best model in explaining performance during the lexical task in terms of both accuracy (Table 5.5) and cognitive load (Table 5.6). Specifically, Model 5 predicted a significant increase in accuracy for participants with higher digit scores performing lexical decisions on nonwords during the suppression and the rhythmic conditions (suppression: b=0.0137, SE=0.0019, t=6.942, p<0.001; rhythm: b=0.0279, SE=0.002, t=13.676, p<0.001). Regarding pupillometry, the model predicted a non-significant increase in pupil diameter for participants with higher digit scores for nonwords during the suppression and the rhythmic conditions (suppression: b=0.0005, SE=0.0019, t=0.272, p=0.786; rhythm: b=-0.0014, SE=0.002, t=-0.702, p<0.483). Overall, higher digit scores predicted a decrease in cognitive load with moderate significance (b=-0.0037, SE=0.0019, t=-1.966, p=0.049). Descriptive statistics and predicted effects are shown in Figure 5.7 and 5.8, respectively.

TABLE 5.5: GLMM comparison for working memory and accuracy in Study 4

Models	Deviance	dAIC	AICw
M.0 Accuracy ∼(1 ∣ id) + (1 ∣ item)	1241	4142	0
M.1 Accuracy ∼condition + (1 ∣ id) + (1 ∣ item)	1234	3618	0
M.2 Accuracy ∼condition + class + (1 ∣ id) + (1 ∣ item)	1225	2884	0
M.3 Accuracy ∼condition * class + (1 ∣ id) + (1 ∣ item)	1202	1001	0
M.4 Accuracy ∼condition * class + digit_span + (1 ∣ id) + (1 ∣ item)	1202	1002	0
M.5 Accuracy ∼condition * class * digit_span + (1 ∣ id) + (1 ∣ item)	1190	0	1

TABLE 5.6: GLMM comparison for working memory and cognitive load in Study 4

Models	Deviance	dAIC	AICw
M.0 Pupil ∼(1 ∣ id) + (1 ∣ item)	-1656	4142	0
M.1 Pupil ∼condition + (1 ∣ id) + (1 ∣ item)	-1657	3618	0
M.2 Pupil ∼condition + class + (1 ∣ id) + (1 ∣ item)	-1657	2884	0
M.3 Pupil ∼condition * class + (1 ∣ id) + (1 ∣ item)	-1657	1001	0
M.4 Pupil ∼condition * class + digit_span + (1 ∣ id) + (1 ∣ item)	-1657	1002	0
M.5 Pupil ∼condition * class * digit_span + (1 ∣ id) + (1 ∣ item)	-1657	0	1

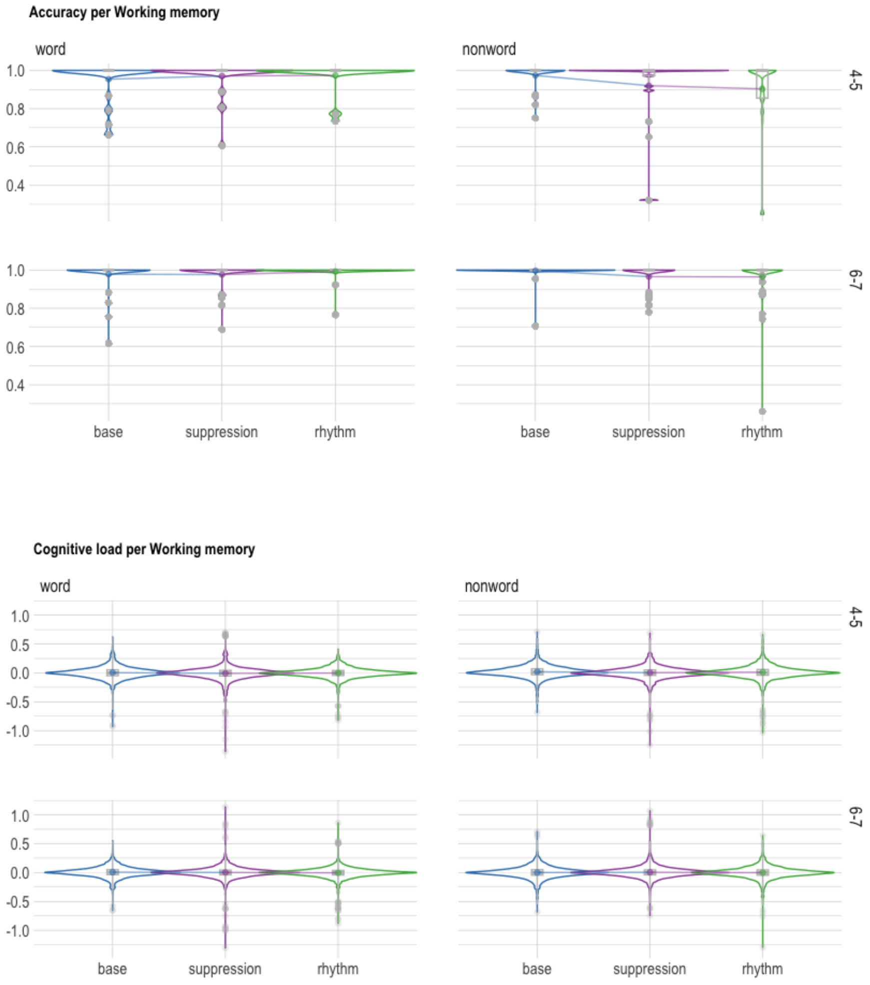


FIGURE 5.7: Descriptive statistics for working memory efficiency. Data are reported for accuracy (top) and baseline corrected pupil (bottom). For this plot, participants were divided in two groups based on their digit span score (4-5 versus 6-7). Data are shown for conditions (i.e., baseline, suppression, and rhythm) and class (i.e., word, non-word).

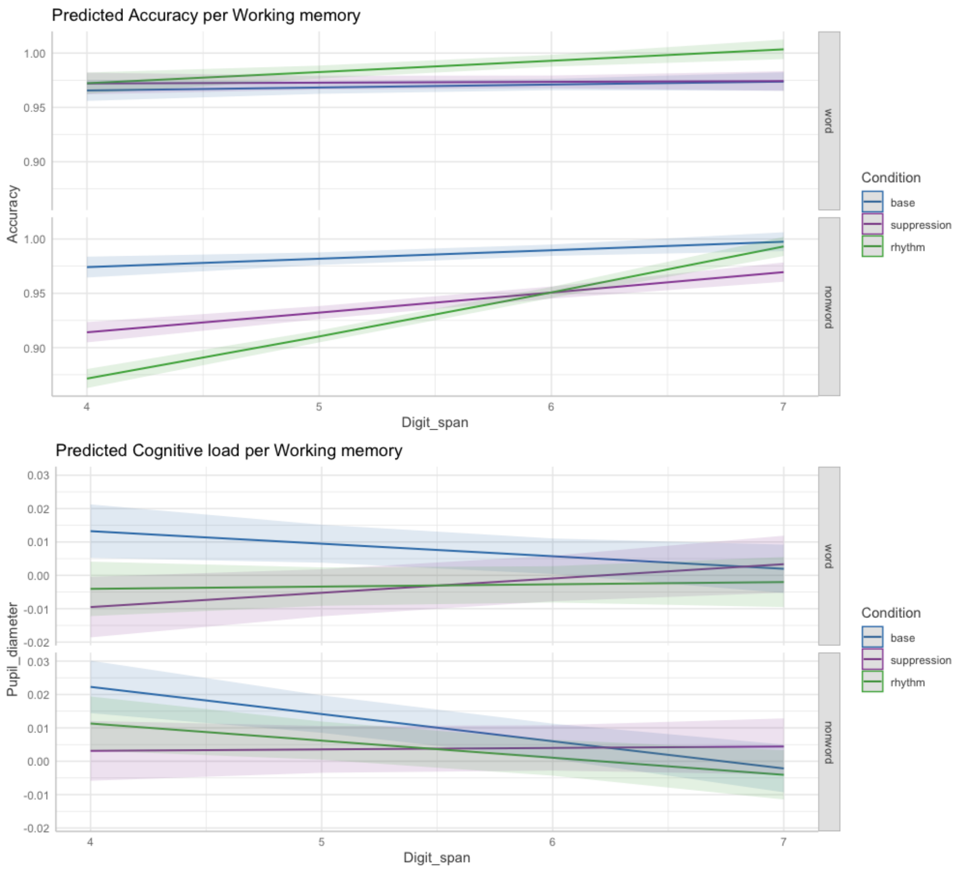


FIGURE 5.8: Effect plots of the best model (M5). The interaction effect of condition (i.e., baseline, suppression, and rhythm), class (i.e., word and non-word), and working memory efficiency (i.e., digit score) is here plotted on predicted values of Accuracy (i.e., the proportion of correct over total answers) and cognitive load (i.e., increase in pupil size under constant luminance).

Predictive effect on reading abilities

Based on the overall model comparison, scores in the self-report checklist for screening literacy difficulties (M4) were the best predictors of reading abilities over music education and working memory, in terms of both accuracy (Table 5.7) and cognitive load (Table 5.8). Refer to Figure 5.4 for visual inspection.

TABLE 5.7: GLMM comparison for lexical accuracy in Study 4

Models	Deviance	dAIC	AICw
M.0 Accuracy ∼(1 | id) + (1 | item)	1241	11380	0
M.1 Accuracy ∼condition + (1 | id) + (1 | item)	1234	10856	0
M.2 Accuracy ∼condition + class + (1 | id) + (1 | item)	1225	10122	0
M.3 Accuracy ∼condition * class + (1 | id) + (1 | item)	1202	8239	0
M.4 Accuracy ∼condition * class * literacy_difficulties + (1 | id) + (1 | item)	1106	0.00	0
M.5 Accuracy ∼condition * class * music_education + (1 | id) + (1 | item)	1180	6456	0
M.6 Accuracy ∼condition * class * working_memory + (1 | id) + (1 | item)	1190	7238	1

TABLE 5.8: GLMM comparison for cognitive load in Study 4

Models	Deviance	dAIC	AICw
M.0 Pupil ∼(1 | id) + (1 | item)	-165646	435.69	0
M.1 Pupil ∼ condition + (1 | id) + (1 | item)	-165715	371.16	0
M.2 Pupil ∼ condition + class + (1 | id) + (1 | item)	-165749	354.41	0
M.3 Pupil ∼ condition * class + (1 | id) + (1 | item)	-165737	339.17	0
M.4 Pupil ∼ condition * class * literacy_difficulties + (1 | id) + (1 | item)	-166104	0.00	0
M.5 Pupil ∼ condition * class * music_education + (1 | id) + (1 | item)	-165784	320.02	0
M.6 Pupil ∼ condition * class * working_memory + (1 | id) + (1 | item)	-165775	329.13	1

5.1.4 Interim discussion

This Study investigated the effect of sensorimotor rhythm on the phonological processing underlying reading abilities in young adults. Specifically, reading abilities were evaluated by means of a lexical decision for word and non-word items (Coltheart et al., 1979; Jacobs and Grainger, 1994). Participants were sixty students in good health, with no sensory or neurological disorders or any familiar language disorder. Therefore, the phonological processing underlying reading in this sample was challenged by including an articulatory suppression to the traditional administration of the lexical task (Martino et al., 2011; Sartori et al., 1995; Colombo et al., 2009). Crucially, participants were assisted by a regular rhythmic cue in performing the lexical task under articulatory suppression in a third, experimental condition. Moreover, the effects of literacy difficulties, previous music education, and working memory efficiency were investigated as predictors of reading performance. Performance was evaluated by behavioral accuracy (i.e., percentage of correct over total responses) and physiological indexes of cognitive load (i.e., increases in pupil diameter).

Overall, self-report difficulties experienced in everyday reading and writing activities (i.e., literacy difficulties) resulted as the best predictor of reading performance in the lexical decision task. Specifically, phonological mechanisms underlying reading in participants with high score of self-report difficulties in literacy predicted to be more easily affected in highly demanding tasks. In fact, if performances among the sample were comparable when processing regular words, the model predicted a significant decrease in accuracy (Figure 5.4) and a significant increase in cognitive load (Figure 5.4) when lexical decisions were performed on non-words under articulatory suppression (i.e., suppression condition). Interestingly, when a regular rhythmic cue was presented (i.e., rhythmic condition) the increase in cognitive load was not observed. Therefore, short-term exposure to a regular rhythmic cue is found to reduce the cognitive load in poor readers performing demanding linguistic tasks. However, the observed reduction of cognitive demand did not result in a boost at the behavioural level. Consistently, accuracy was comparable between suppression and rhythmic conditions in processing non-words.

The effects of long-term exposure to rhythm were then explored by including years of musical education as a predictor of reading abilities across conditions. Results showed that individuals who received long-term exposure to rhythm through active music practice are predicted to benefit the most from a regular rhythmic cue in performing a reading task. Consistently, participants with 5 to 10 years of music practice were found to perform lexical decisions on words and non-words with significantly higher accuracy. Furthermore, significantly lower levels of cognitive demand were required to process words and non-words for participants with 1 to 10 years of music education. Lastly, the effect of working memory efficiency was estimated resulting as a significant predictor of performances across conditions. Specifically, higher scores in the digit span test predicted higher levels of accuracy (especially in the rhythmic condition) and lower cognitive load across conditions. Therefore, results go in the direction of a positive effect of short-term and long-term exposure to rhythm in the underlying mechanisms supporting reading abilities in young adults.

Specifically, rhythm seems to play a positive effect at the level of phonological processing. This interpretation is based on the experimental manipulation applied to this Study. Specifically, articulatory suppression was included to challenge readers with an interfering effect produced on the phonological loop (Baddeley, 1993). In fact, the phonological loop is the working space where verbal information is encoded and processed to formulate an output. The continued articulation of irrelevant information can thus disrupt the correct encoding and processing in reading, by overloading the underlying phonological mechanism (Baddeley, 1993). Crucially, the effect of rhythm in this Study was observed only with respect to the articulatory suppression condition. Therefore, it has been interpreted as an aid in reducing the interference of articulatory suppression on the phonological loop. The involvement of rhythm in phonological processes underlying reading might be further confirmed by the associations observed between working memory span and performances in the rhythmic condition. Accordingly, working memory is a required component for operating on the material encoded in the phonological store (Baddeley, 1993).

Therefore, to successfully perform a lexical decision task, verbal information must be encoded, temporarily stored, and then processed to produce an output. Although articulatory suppression is known to impact this process chain by interfering in phonological processes, here rhythm was found to reduce this effect, with better chances for individuals with higher working memory span. A further step forward for a deeper knowledge of the mechanisms underlying this effect is represented by pupillometry. In fact, combining the behavioural measures with physiological indexes might reveal further information on the beneficial effect of rhythm. Consistently, this Study seems to suggest that the facilitating effect of rhythm on phonological mechanisms in reading might involve the cognitive load underlying the process. Specifically, when rhythm is found to predict better performances in accuracy, a decrease in cognitive load is observed. Vice versa, when verbal information requires higher attentional demands to be processed (e.g., suppression condition or non-words) an increase in cognitive load was registered together with a decrease in accuracy. Accordingly, pupil diameter is known to be a reliable index of the cognitive load underlying task processing in the literature (Sirois and Brisson, 2014). In the context of this Study, the facilitating effect of rhythm on the phonological processing underlying reading might thus be interpreted in terms of cognitive load. Specifically, a regular rhythm might cue participants in performing the lexical task underlying articulatory suppression. Therefore, a regular rhythmic cue could contribute in reducing the cognitive demand required by the articulatory suppression thus releasing resources for phonological and working memory processes. Interestingly, the reduced cognitive load registered in the rhythm condition was not associated with increased accuracy in performance for poor readers. In fact, reading difficulties were found to significantly predict a fall in processing non-words under articulatory suppression but no improvements were observed when the rhythmic cue was introduced.

This result might be interpreted as an empirical support of the atypical rhythm processing found among populations suffering from language impairments (Ladányi et al., 2020; Lense et al., 2021; Goswami, 2011). Accordingly, individuals diagnosed with or at risk for dyslexia were found to show difficulties in a vast range of rhythmic skills including atypical entrainment to external rhythms (Thomson et al., 2006; Thomson and Goswami, 2008; Ladányi et al., 2020). However, individuals with dyslexia were also found to improve in linguistic tasks after short-term exposure to a regular rhythmic cue (Canette et al., 2019, 2020; Przybylski et al., 2013). Therefore, targeting the atypical rhythmic skills observed in dyslexia might produce positive cascading effects on linguistic abilities. Accordingly, the benefits resulting from short and long-term training involving rhythm are well known in the field (Bonacina et al., 2020; Reifinger Jr, 2019). However, mixed or opposite results are also reported in the literature (Gordon et al., 2015, for a meta-analysis). Therefore, further research is needed to weigh the effect of short- and long-term rhythmic exposure on linguistic abilities in poor readers. However, the results reported in this Study might be unlikely explained by a disadvantage of the rhythmic cue in processing language. In fact, the rhythmic condition overall significantly predicted better performance in terms of accuracy and decreasing cognitive load. Therefore, other possible explanations might be advanced. For instance, poor readers are known to show impairments in processing rapid sequential stimuli (Hari and Renvall, 2001; Tallal, 2004). The 2 Hz frequency could thus have turned out to be too fast with respect to the processing time of poor readers. Furthermore, impaired motor entrainment at 2 Hz is reported in dyslexia (Thomson et al., 2006; Thomson and Goswami, 2008). Although no significant difference was found for spontaneous motor tempo, adults with dyslexia were found to significantly differ from controls in synchronizing to 550 ms IOI sequences (1.81 Hz) in a recent study by Canette et al. (2020). Therefore, it remains an open question whether poor readers in the present Study did not show any benefit from the rhythmic cue because of a general impairment in processing rhythm or whether the decrease in cognitive load might result in better performances at the behavioural level with slow tempo rate. Preliminary studies combining traditional interventions for language and reading with an interactive metronome suggest that results might go in this direction. However, further research is needed to confirm these predictions.

5.1.5 Conclusions to Study 4

The present Study aimed at exploring whether sensorimotor entrainment to rhythm could boost the phonological processes underlying reading in young adults. Specifically, reading abilities were evaluated in a sample of sixty individuals from good to at risk readers, based on their scores in a self-report screening for dyslexia. A regular rhythmic cue significantly predicted a decrease in cognitive load while reading. However, the facilitating effect of rhythm resulted in increased accuracy for proficient readers only. Results were explained in terms of perceptual and cognitive mechanisms underlying different levels of reading proficiency. Specifically, the effect of rhythm was interpreted taking into account the role of phonological processing, working memory, cognitive load, and processing speed. A possible explanation related to the impairments in rapid stimulus sequence (Hari and Renvall, 2001, RSS) was proposed to account for the lack of transition from the facilitating effect observed on cognitive load to improved accuracy in poor readers. Moreover, a further investigation into the effect of spontaneous motor tempo and interactive metronome was proposed (Ritter et al., 2013). Within the TSF (Goswami, 2011), the facilitating effect on cognitive load was explained in terms of neural entrainment eliciting Delta-band processes with cascading effects on attention, working memory, and phonological processes (Thomson et al., 2006; Thomson and Goswami, 2008). Specifically, the sensorimotor entrainment of vocal production to a regular rhythm is interpreted as supportive in containing the effect of articulatory suppression on phonological mechanisms. Accordingly, sensorimotor coupling is considered one of the three main processes underlying the general rhythmic abilities that support language across development, together with fine auditory processing and neural entrainment (Fiveash et al., 2021; Nayak et al., 2022; Ladányi et al., 2020).

Consistently, years of active practice with a musical instrument significantly predicted lower cognitive load and increasing accuracy in reading. Moreover, participants with higher scores of music education were found to benefit the most from the rhythmic cue, confirming the interconnected efficacy of short- and long-term exposure to rhythm in literacy development (Schön and Tillmann, 2015; Miendlarzewska and Trost, 2014; Russo and Valenza, 2021). In the cross-domain transfer from music to language, synchronizing motions to an external rhythm might thus elicit neural entrainment supporting attention, working memory, and phonological processing in literacy skills. This transfer of effects through sensorimotor coupling might be explained by referring to the neural networks involved in processing rhythm across music and language in high vocal learner species (Patel, 2021, see Chapter 2). Accordingly, the positive effects of rhythm observed in perceiving, producing, and reading language might be maximized in writing, given the motor-related nature of this advanced literacy skill. To test this hypothesis, participants were called to perform a second experimental section involving the sensorimotor effects of rhythm in a writing task, illustrated and discussed in the next section.

5.2 Study 5: influence of sensorimotor rhythm on writing abilities

The evolution of writing – a system of graphic marks conveying meaning – is probably one of the most relevant achievements in human history. From ideographic products, throughout the first cuneiform scripts, the generative and creative code involving modern alphabetic systems has not only determined the spread of knowledge around cultures and throughout history. In fact, it has also shaped the development of information processing in the cognitive system to operate on and transmit large amounts of information through abstraction (Schmandt-Besserat, 2014). Consistently, the ontogenetic development of writing abilities does not require linguistic skills only; by contrast, mastering the gesture of handwriting involves also general cognitive and fine motor abilities, as well as the complex interplay between these two. Accordingly, the development of writing abilities builds on phoneme-grapheme knowledge, fine kinematic processes, and the hierarchical planning of each unit through time and space (Lashley et al., 1951; Fitch, 2014). Consequently, underlying possible impairments have been identified across atypical trajectories leading to dysgraphia and dyslexia, with considerable overlapping and comorbidity between these two (Döhla et al., 2018; McCloskey and Rapp, 2017). Among others, difficulties in the orienting of attention (Facoetti et al., 2003), rapid sensory processing (Tallal, 1980; Stein and Walsh, 1997), motor control as well as automatization and motor programming (Cheng-Lai et al., 2013; Lam, 2011; Nicolson and Fawcett, 1990) are reported in the literature. To draw these findings within an integrated framework, a general timing deficit has been proposed to underlie the perceptual and phonological difficulties experienced in dyslexia (Goswami, 2011, TSF). Accordingly, individuals with dyslexia are found to perform worse than controls in non-linguistic timing tasks, including tapping to a metronome, reproducing rhythms, and correctly encoding motor and auditory rhythms (Wolff, 2002; Thomson et al., 2006; Thomson and Goswami, 2008). The role of general rhythmic abilities has been further highlighted by recently formulated frameworks on language development, as illustrated in previous Chapters (Ladányi et al., 2020, ARRH; Fiveash et al., 2021, PRISM; Nayak et al., 2022, MAPLE).

These frameworks collected evidence on several rhythmic deficits spanning across language and language-related disorders, including developmental dyslexia, developmental language disorder, stuttering, developmental coordination disorder, and attention deficit hyperactivity disorder (Colling et al., 2017; Bedoin et al., 2016; Chang et al., 2016; Puyjarinet et al., 2017; Carrer, 2015). However, difficulties regarding writing abilities have been significantly less investigated. Given the commonalities and the high rate of comorbidity between developmental dyslexia and dysgraphia, a general timing deficit can be hypothesized underlying writing deficit in dysgraphia. A few studies so far investigated this hypothesis by testing the rhythmic structure of motor timing planning in handwriting (Pagliarini et al., 2017, 2015).

Accordingly, most of the research on dysgraphia has focused on the kinematic aspects of handwriting (Accardo et al., 2013; Di Brina et al., 2008; Kushki et al., 2011). Differently, the two studies from Pagliarini et al. (2017, 2015) originally investigated the adherence of handwriting to two general principles of a rhythmic organization in human movement: isochrony and homothety. Specifically, homothety (Lacquaniti et al., 1983) predicts that the relative duration of the individual letters composing a word remain constant across changes in the overall duration whereas isochrony (Viviani and Terzuolo, 1982) establishes that the speed of movement execution is proportional to the length of the movement trajectory. Therefore, both principles require fine timing abilities to form an internal temporal structure governing movements, thus determining the rhythmic dimension of handwriting (Pagliarini et al., 2017). Consistently, individuals are found to comply with both principles from the first grade of primary school, thus suggesting the existence of invariants governing the internal representation of rhythm in handwriting (Pagliarini et al., 2017). Interestingly, children with developmental dyslexia, with and without dysgraphia, were found to fail to adhere to the rhythmic principles of isochrony and homothety (Pagliarini et al., 2015). Therefore, rhythm might play a significant role in the typical development of writing skills by guiding the temporal organization of motor executions.

However, further research is needed to extend these findings. For instance, less is known about the short- and long-term effect of rhythm on the efficiency of writing abilities in adults. Moreover, despite rhythm processing has been suggested to underlying both reading and writing abilities, further research is needed to strengthen these findings and clarify the possible mechanisms through which it can lead to possible dysfunction in writing abilities. Moreover, understanding the effects of rhythm on writing might help in identifying the preserved abilities of children and adults with linguistic impairments thus offering opportunities for targeted training and interventions. Therefore, the present Study aims at investigating the effect of short- and long-term rhythm on the cognitive and motor processes underlying writing abilities.

5.2.1 Aim and research questions

The present Study aims at exploring the beneficial effect of rhythm on writing abilities in young adults. Specifically, handwriting quality and kinematics will be investigated in the sample of sixty young adults participating in Study 4 (section 5.1), called to perform a second experimental section consisting of a dictation task. This task - adapted from Colombo et al. (2009) - consisted of a list of words auditory presented to and written by participants. Performances will be evaluated in terms of product quality (i.e., legibility), kinematic indexes (i.e., average speed and pressure of stroke), and adherence to the isochrony principle (i.e., the ratio between speed and length of movements). As in Study 4, the dictation task was performed in a baseline condition, under articulatory suppression, and in presence of a rhythmic cue. Therefore, possible difficulties in the phonological mechanisms underlying writing skills will be investigated by means of articulatory suppression, known to interfere with verbal manipulation in the phonological loop (Baddeley, 1993; Colombo et al., 2009; Martino et al., 2011). Consistently, possible difficulties are expected to emerge in the phonological processing underlying writing abilities when performing the dictation task under articulatory suppression. Moreover, given the overlapping between the dyslexic and dysgraphic profiles reported in previous literature, a decrease in performance under articulatory suppression is expected for participants with lower reading abilities. Specifically, reading proficiency was evaluated by the self-report screening for dyslexia completed by participants prior to the first experimental section (Vinegard, 1994, see Study 4 in section 5.1). Furthermore, the role of working memory efficiency in contributing to the phonological processes underlying writing will be explored by investigating the modulating effect of digit span levels on handwriting. Lastly, the beneficial effect of rhythm on writing will be investigated by means of a regular rhythmic cue added in a third experimental condition in which participants performed under articulatory suppression but in presence of a regular rhythm cue. Specifically, paralleling the findings from speech and reading (Brandt et al., 2012; Ozernov-Palchik and Patel, 2018), a regular rhythmic cue is also expected to improve writing skills.

Furthermore, the core involvement of sensorimotor coupling as underlying mechanism in processing rhythm and language (Patel and Iversen, 2014; Fiveash et al., 2021) is here hypothesized to elicit an even more effective cue, compared to reading performances, given the intrinsic motor nature of writing. Therefore, the interfering effect of suppression is expected to be significantly reduced by cueing vocal articulation and motor production with a regular external rhythm. According to the TSF (Goswami, 2011), a regular rhythm might elicit Delta processes in neural oscillations resulting in better attention allocation and processing. Therefore, better performance is expected in performing the dictation task in the rhythmic condition. Furthermore, the effect of a long-term exposure to rhythm (measured in years of received musical education; see Study 4) will be evaluated as a potential predictor of writing proficiency. Specifically, having received a musical education is expected to result in better writing performance thus extending previous findings on the link between music and literacy development (Flaugnacco et al., 2015; Overy, 2003; Cogo-Moreira et al., 2013).

5.2.2 Method

Participants

Participants from Study 4 participated in this Study as a second experimental section to the whole project. Therefore, please refer to section 5.2.2 for further detail on this sample. Participants signed a consent form before starting the first experimental section.

Stimuli and Materials Words for the dictation task consisted of 48 items from Colombo et al. (2009). Words were divided into three lists of 16 items each, balanced for syllable number and frequency. The digit span test, self-report difficulties in everyday literacy abilities, and music education were investigated in prior to the first experimental section. Therefore, please refer to section 5.2.2 for further details on the stimuli and materials by which these measures were collected.

Apparatus

Kinematic measures were collected by adapting the dictation task (Colombo et al., 2009, i.e., paper-and-pencil) to a digitalized version. Specifically, the task was performed by means of an iPad Air of fourth generation and an Apple Pencil of second generation. The iPad has a 10.9-inch (diagonal) LED-backlit multi-touch display and a resolution of 2360x1640 pixels at 264 ppi (pixels per inch). The Apple Pencil is sensitive to tilt and pressure and, while using it, it is possible to place a hand on the iPad screen as with paper-and-pencil. Participants were instructed to write on a white screen positioned with the longer side in the horizontal position while an algorithm collected data on speed, pressure, and length of strokes. Rhythmic stimulation was provided via a metronome paced at 2 Hz (120 bpm). To balance the presence of an auditory stimulation across conditions, white noise was played in the background.

Experimental design and procedure

Participants completed the online questionnaire prior to the first experimental session (see section 5.1.2). The writing session was carried out at the eye-tracker laboratory at the Department of Developmental Psychology and Socialization of the University of Padua. After signing the consent form, participants were presented with the dictation task. Firstly, the writing session began with two training trials presented to familiarize participants with the task and instruments. The experimental design consisted of three conditions: i) a baseline condition, in which the dictation task was performed with no interference or cue; ii) a suppression condition, in which participants were instructed to perform the dictation task continuously repeating the syllable *La*; and iii) a rhythmic condition, in which participants performed the dictation task under articulatory suppression cued by a metronome. In each condition, one of the three lists of items was presented. As for Study 4, the experimental design was counterbalanced based on i) order presentation, and ii) list of items (see Section 5.1.2). During the experimental section, each participant was tested individually, seating at a desk in a quiet room.

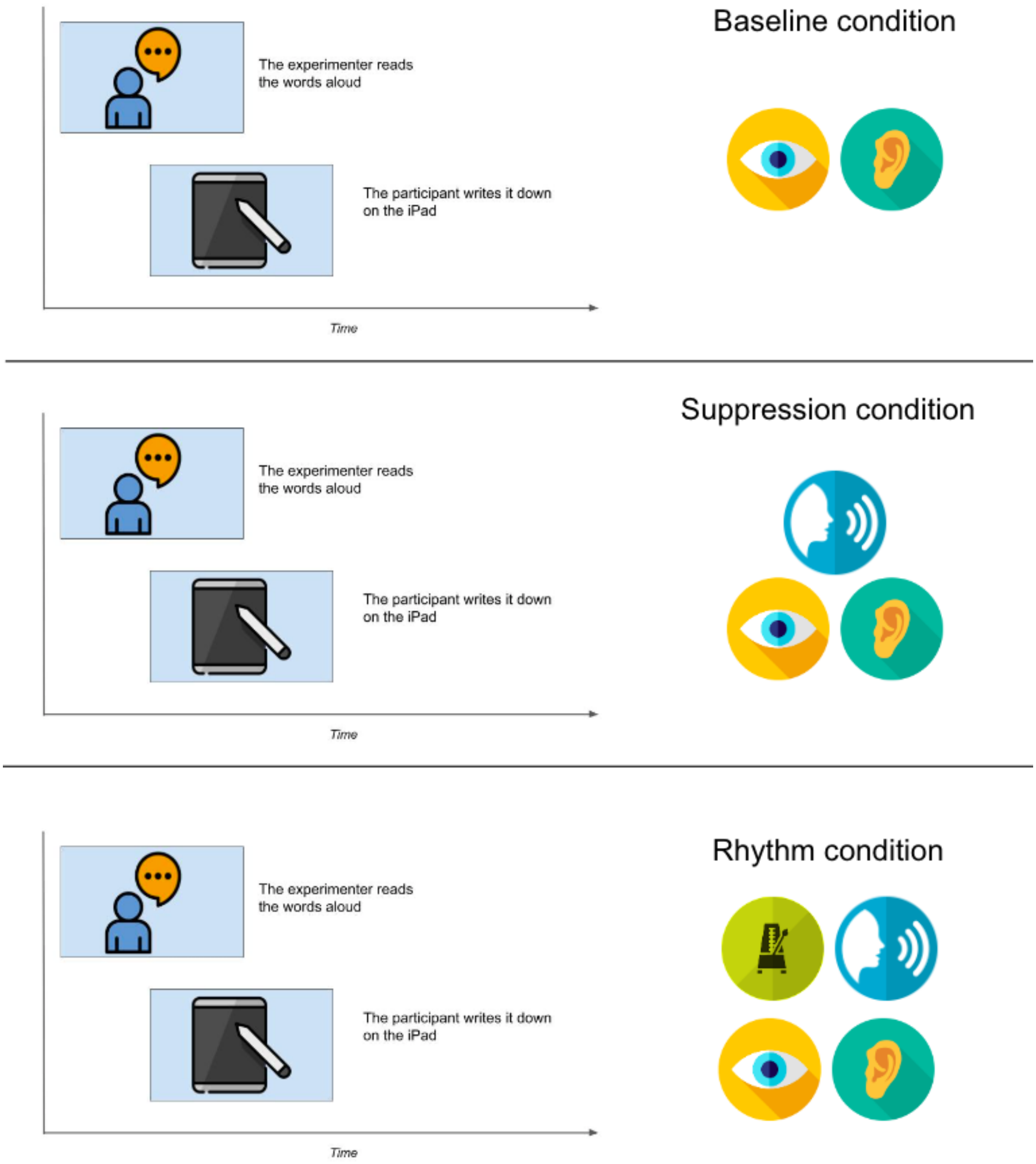


FIGURE 5.9: Experimental design from Study 5. For each condition, participants were instructed to perform the dictation task (left). Across each condition, the experimenter read the list word by word while participants transcribed items on the iPad. Therefore, the baseline condition mainly interested the visual and auditory sensory modality. In the suppression condition, participants were instructed to perform the task under articulatory suppression, thus vocally producing a verbal output (i.e., *La*) while writing words. In the rhythmic condition, a regular rhythmic cue (i.e., a metronome paced at 2 Hz) assisted participants in performing the dictation task under articulatory suppression.

Statistical analyses

Performances on the dictation task were coded in terms of product quality, kinematic indexes, and isochrony. Specifically, product quality was firstly assessed in terms of legibility, through a Likert scale, by three independent evaluators blinded to hypothesis and design (Barnett et al., 2018). Kinematic indexes included average speed and average pressure of stroke (Rosenblum et al., 2003; Kushki et al., 2011). Lastly, isochrony was computed as the ratio between average speed and stroke length (Pagliarini et al., 2015, 2017). Working memory scores were obtained from the digit span test performed in the first experimental session (Hartman, 2009; Greiffenstein et al., 1994). Difficulties in literacy abilities were coded from the self-report screening for dyslexia completed in the first experimental session (Vinegard, 1994, see Section 5.2.2). Specifically, participants with a total score above the third quartile of the data distribution were coded as low-proficient (n=12) whereas the rest of the sample was coded as with high-proficient (n=48). Participants who didn't receive any prior music education were coded as non-musicians (n=39) whereas participants who received from 1 to 10 years of music education were coded as musicians (Zhang et al., 2020, n=21). Product quality, kinematic indexes, and isochrony were modeled as continuous variables with gamma and gaussian distributions, by means of Generalized Mixed-Effects models (Nieuwenhuis et al., 2012; Fox, 2015). Fixed effects were gradually added to the null models accounting for participant and item as random effects. To investigate the predictive value of each effect, a model comparison was performed using the likelihood ratio test (LTR), Akaike Information Criterion (Akaike, 1974, AIC), and AIC weight as indexes of the goodness of fit.

5.2.3 Results

The effect of literacy proficiency, music education, and working memory on product quality (section 5.2.3), kinematics (section 5.2.3), and rhythmic organization (section 5.2.3) of handwriting across conditions is presented as follows.

Product Quality

From the model comparison, Model 3 resulted as the best model in explaining the effect of literacy proficiency on product quality across conditions (Table 5.9). Specifically, M3 predicted lower quality scores for low-proficient readers (b=-0.02665, SE=0.00294, t=-9.054, p<0.001). Moreover, scores were predicted to significantly increase in the rhythm condition for low-proficient readers (b=0.0238, SE=0.0009, t= 24.294, p<0.001). Investigating the modulating effect of music education, Model 3 resulted as the best model accounting for the interaction between years of music education and conditions (Table 5.10). Overall, music education resulted as a significant predictor of an increase in the quality score (b=0.0119, SE=0.0046, t=2.575, p=0.010). Specifically, the strongest effect of music education was observed in containing articulatory interference in the suppression condition (b=0.01504, SE=0.00156, t=9.599; p<0.001). Lastly, the effect of working memory was better explained by Model 3 accounting for the interaction between digit span and conditions (Table 5.11). Overall, higher levels working memory predicted a general increase in product quality (b=0.04516, SE=0.01169, t=3.862, p<0.001). Moreover, participants with lower levels of working memory performed significantly better in the rhythmic compared to the baseline and suppression conditions (b=-0.0314, SE=0.00394, t=-7.991, p<0.001). Among predictors, differences in product quality across conditions were better predicted by literacy proficiency level (Table 5.12). Descriptive statistics and model effect plots are presented in Figure 5.10, Figure 5.11, and Figure 5.12.

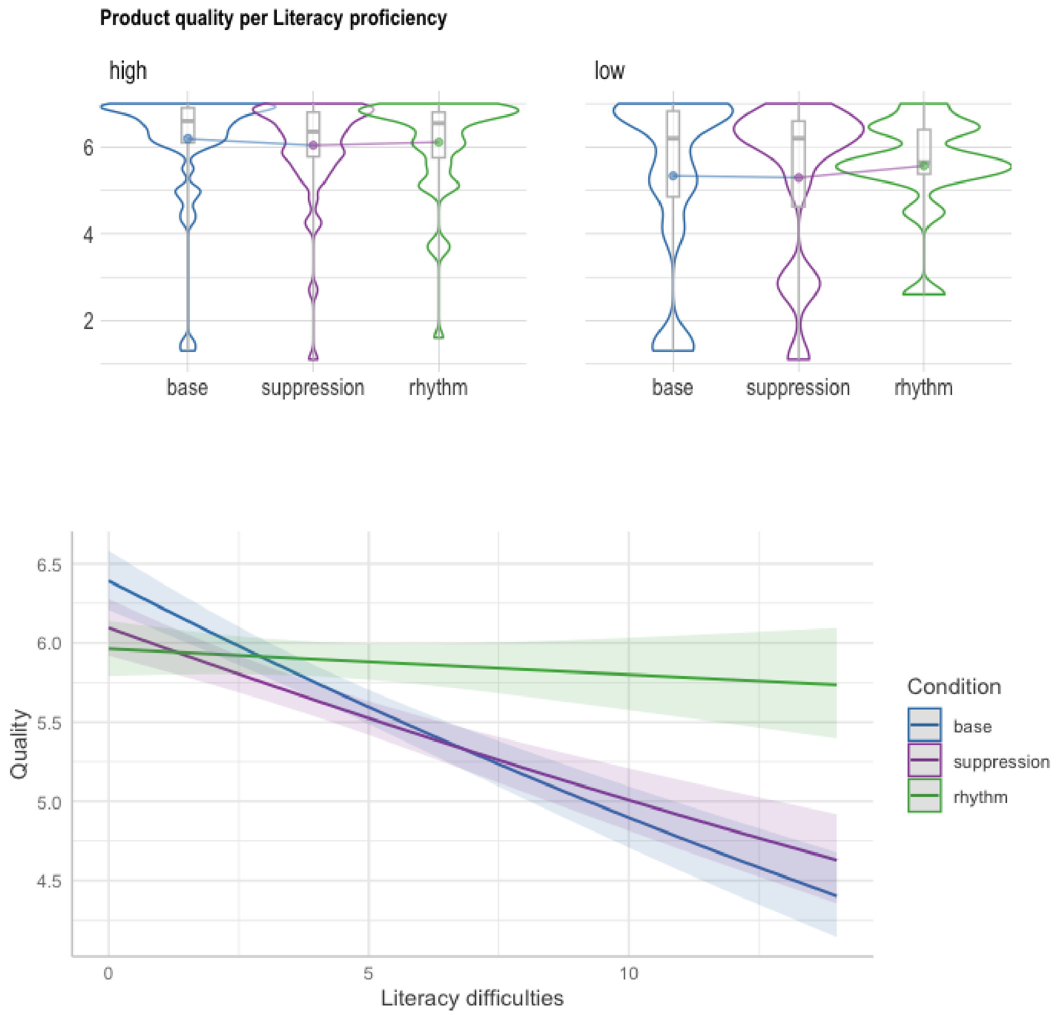


FIGURE 5.10: Descriptive statistics and effect plot of product quality per literacy proficiency. Data distribution with means of quality scores per condition and literacy proficiency (top). Effect plot of M3 showing predicted values of quality score in conditions as a function of literacy difficulties (bottom).

TABLE 5.9: GLMM comparison for literacy and quality in Study 5

Models	Deviance	dAIC	AICw
M.0 Quality ∼(1 ∣ id) + (1 ∣ trial)	141.1363	811.21	0
M.1 Quality ∼condition + (1 ∣ id) + (1 ∣ trial)	138.0011	613.28	0
M.2 Quality ∼condition + literacy proficiency + (1 ∣ id) + (1 ∣ trial)	138.0026	614.21	0
M.3 Quality ∼condition * literacy proficiency + (1 ∣ id) + (1 ∣ trial)	128.8371	0.00	1

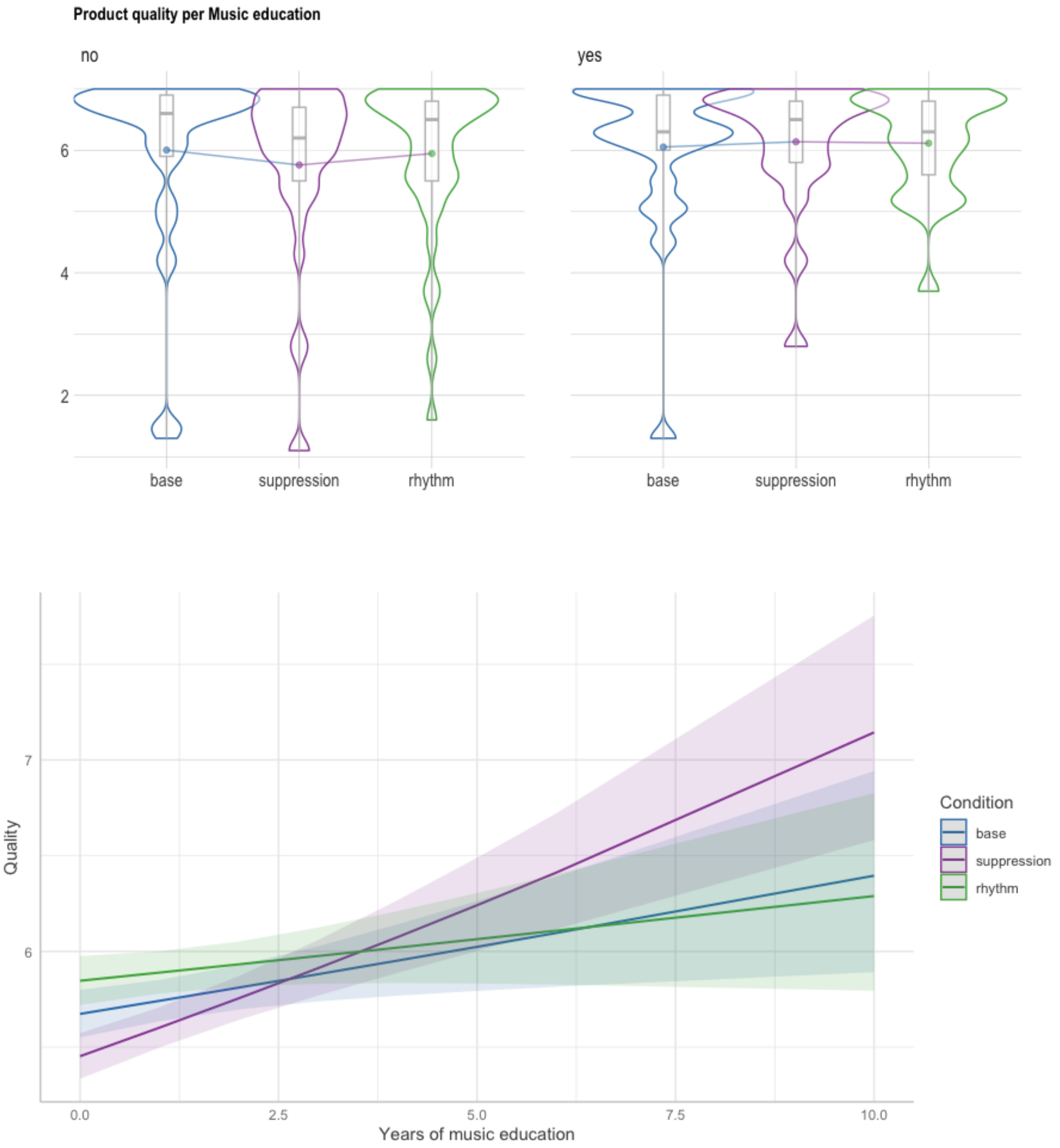


FIGURE 5.11: Descriptive statistics and effect plot of product quality per music education. Data distribution with means of quality scores per condition and music education (top). Effect plot of M3 showing predicted values of quality score in conditions as a function of years of music education (bottom).

TABLE 5.10: GLMM comparison for music education and quality in Study 5

Models	Deviance	dAIC	AICw
M.0 Quality ∼(1 ∣ id) + (1 ∣ trial)	141.1363	261.55	0
M.1 Quality ∼condition + (1 ∣ id) + (1 ∣ trial)	138.0011	63.62	0
M.2 Quality ∼condition + music education + (1 ∣ id) + (1 ∣ trial)	138.0017	65.20	0
M.3 Quality ∼condition * music education + (1 ∣ id) + (1 ∣ trial)	136.9401	0.00	1

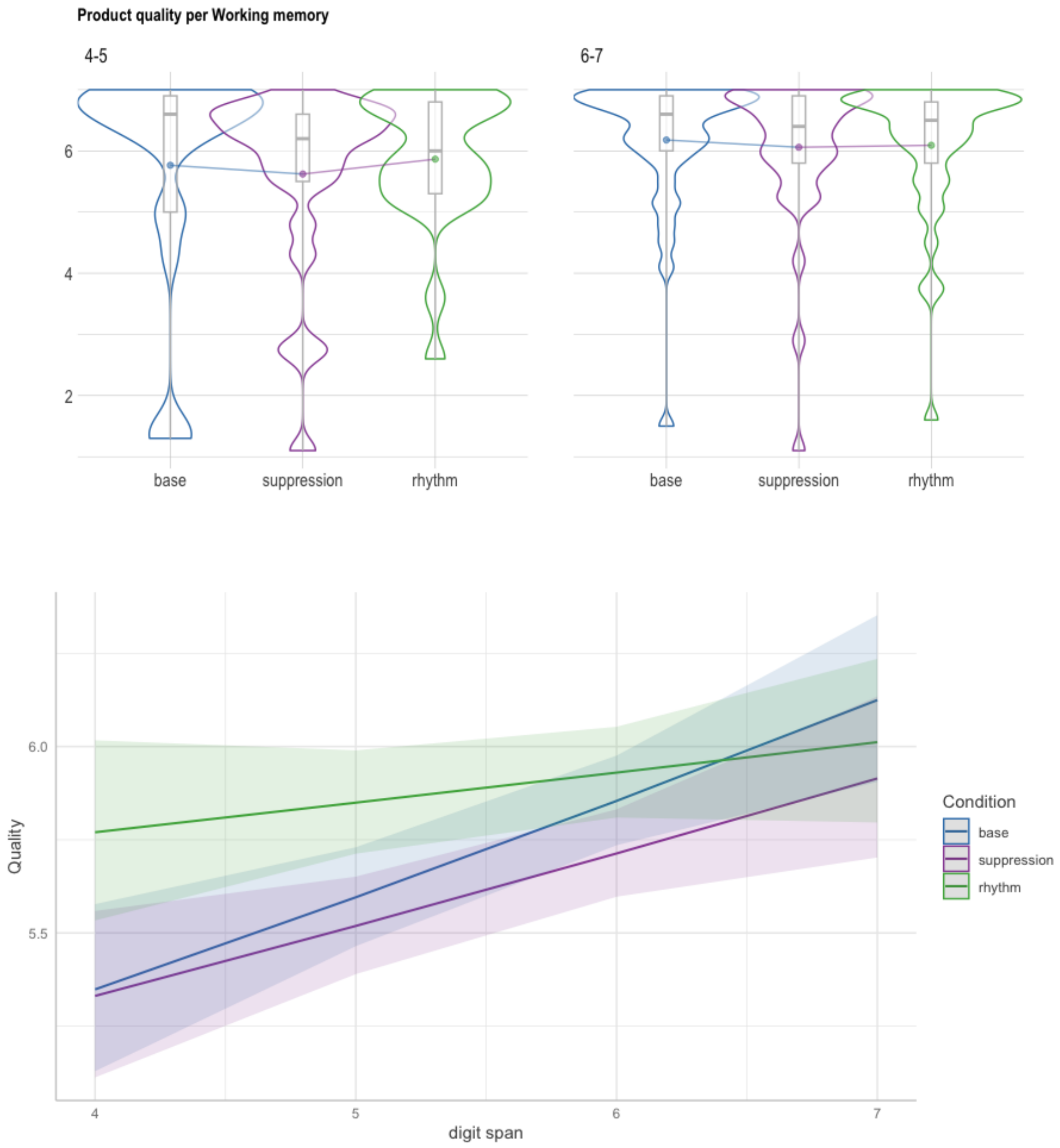


FIGURE 5.12: Descriptive statistics and effect plot of product quality per working memory. Data distribution with means of quality scores per condition and digit span score (top). Effect plot of M3 showing predicted values of quality score in conditions as a function of digit span score (bottom).

TABLE 5.11: GLMM comparison for working memory and quality in Study 5

Models	Deviance	dAIC	AICw
M.0 Quality ∼(1 ǀ id) + (1 ǀ trial)	141.1363	261.55	0
M.1 Quality ∼condition + (1 ǀ id) + (1 ǀ trial)	138.0011	63.62	0
M.2 Quality ∼condition + working memory + (1 ǀ id) + (1 ǀ trial)	138.0015	65.36	0
M.3 Quality ∼condition * working memory + (1 ǀ id) + (1 ǀ trial)	136.9401	0.00	1

TABLE 5.12: GLMM comparison for product quality in Study 5

Models	Deviance	dAIC	AICw
M.1 Quality $\sim$condition * literacy proficiency + (1 | id) + (1 | trial)	128.8371	0	1
M.2 Quality $\sim$condition * music education + (1 | id) + (1 | trial)	135.298	440.82	0
M.3 Quality $\sim$condition * working memory + (1 | id) + (1 | trial)	136.9401	549.66	0

Kinematic indexes

Kinematic indexes include average speed and pressure. Specifically, a significant decrease in average speed was predicted by low literacy in the suppression (b= -0.0005, SE=0.00007, t=-6.951, p<0.001) but not in the rhythmic condition (b=-0.0001, SE=0.00007, t=-1.711, p=0.087; M3, Table 5.13). Music education did not contribute any significant effect to the models (b=0.0007, SE=0.0007; t=1.001, p=0.317). By contrast, higher working memory scores predicted an overall increase in average speed across conditions (b=0.0040, SE=0.0018, t=2.126, p=0.033; M2, Table 5.14). Moreover, higher working memory scores significantly predicted an increase in average pressure for the suppression condition (b=0.0539, SE=0.0045, t=11.889, p<0.001) and a significant decrease in the rhythmic condition (b=0.0347, SE=0.0044, t=7.865, p<0.001; M3, Table 5.15). Descriptive statistics and effect plots are presented in Figure 5.13 and Figure 5.14.

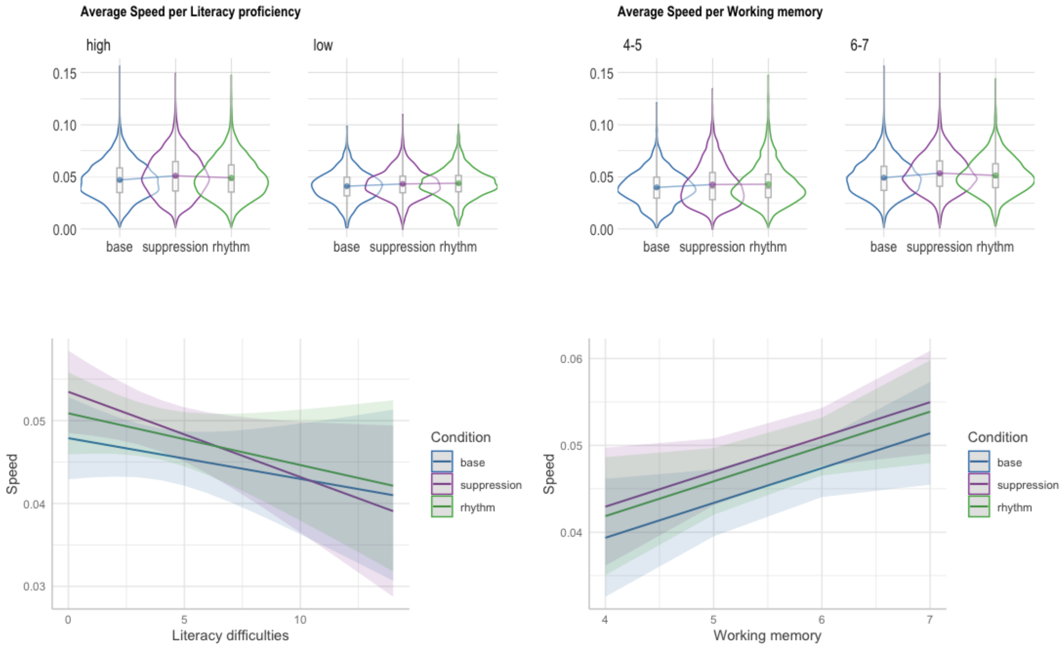


FIGURE 5.13: Descriptive statistics and effect plot of average speed per literacy proficiency and working memory. Data distribution with means of average speed (top) per literacy proficiency (left) and working memory (right). Effect plot of M3 showing predicted values of quality score in conditions as a function of literacy difficulties (bottom, left) and working memory (bottom, right).

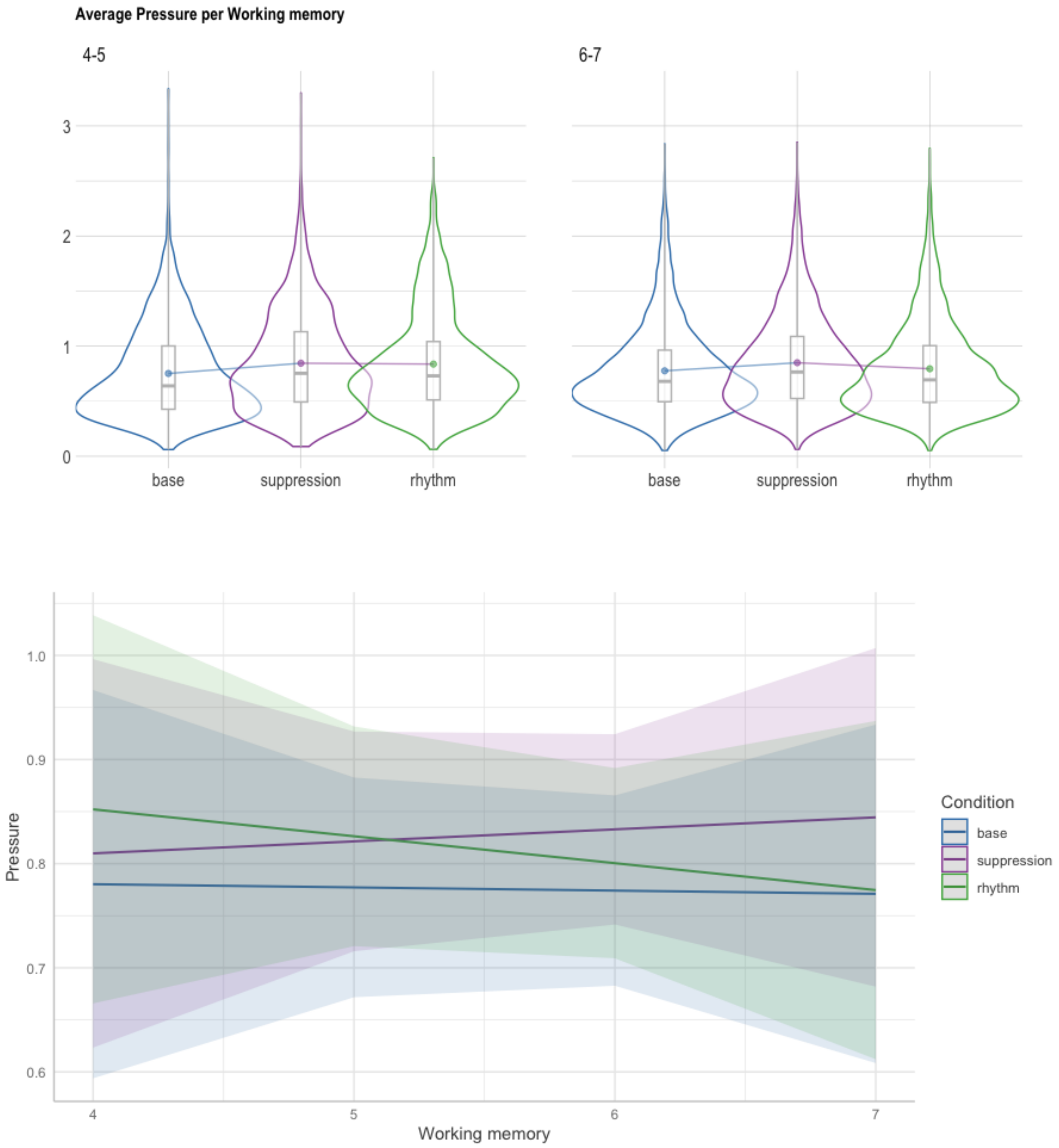


FIGURE 5.14: Descriptive statistics and effect plot of average pressure per working memory. Data distribution with means of average pressure per working memory (top). Effect plot of M3 showing predicted values of quality score in conditions as a function of working memory (bottom).

TABLE 5.13: GLMM comparison for literacy and speed in Study 5

Models	Deviance	dAIC	AICw
M.0 SpeedAvg ∼(1 ∣ id) + (1 ∣ trial)	-129728	225.16	0
M.1 SpeedAvg ∼Literacy proficiency + (1 ∣ id) + (1 ∣ trial)	-129717	238.41	0
M.2 SpeedAvg ∼Literacy proficiency + condition + (1 ∣ id) + (1 ∣ trial)	-129945	14.31	0
M.3 SpeedAvg ∼Literacy proficiency * condition + (1 ∣ id) + (1 ∣ trial)	-129964	0.00	1

TABLE 5.14: GLMM comparison for working memory and speed in Study 5

Models	Deviance	dAIC	AICw
M.0 SpeedAvg ∼(1 ∣ id) + (1 ∣ trial)	-129728	215.92	0
M.1 SpeedAvg ∼Working memory + (1 ∣ id) + (1 ∣ trial)	-129722	224.24	0
M.2 SpeedAvg ∼Working memory + condition + (1 ∣ id) + (1 ∣ trial)	-129950	0.00	1
M.3 SpeedAvg ∼Working memory * condition + (1 ∣ id) + (1 ∣ trial)	-129928	26.27	0

TABLE 5.15: GLMM comparison for working memory and pressure in Study 5

Models	Deviance	dAIC	AICw
M.0 PressureAvg ∼(1 ∣ id)	5427	142.53	0
M.1 PressureAvg ∼Working memory + (1 ∣ id)	5431	148.59	0
M.2 PressureAvg ∼Working memory + condition + (1 ∣ id)	5304	25.62	0
M.3 PressureAvg ∼Working memory * condition + (1 ∣ id)	5274	0.00	1

Isochrony

A lower adherence to isochrony was significantly predicted by high levels of dyslexia risk (b=-0.00005, SE=0.00001, t=-4.093, **??**<0.001; M4, Table 5.16) and low levels of working memory (b=0.00007, SE=0.00001, t=6.145, **??**<0.001; M4 Table 5.17). By contrast, music education predicted a general increase in velocity (b=0.00594, SE=0.00297, t=1.998, **??**=0.046; M4, Table 5.18). Overall, working memory resulted as the best predictor (M3, Table 5.19). Descriptive statistics and effect plots are presented in Figure 5.15, Figure 5.16, and Figure 5.17.

TABLE 5.16: GLMM comparison for literacy and isochrony in Study 5

Models	Deviance	dAIC	AICw
M.0 SpeedAvg ∼(1 ∣ id)	-129728	2339.87	0
M.1 SpeedAvg ∼Length + (1 ∣ id)	-131824	246.61	0
M.2 SpeedAvg ∼Length + literacy proficiency + (1 ∣ id)	-131818	254.31	0
M.3 SpeedAvg ∼Length * literacy proficiency + (1 ∣ id)	-131813	261.63	0
M.4 SpeedAvg ∼Length * literacy proficiency + condition + (1 ∣ id)	-132078	0.00	1
M.5 SpeedAvg ∼Length * literacy proficiency*condition + (1 ∣ id)	-132028	62.34	0

TABLE 5.17: GLMM comparison for music education and isochrony in Study 5

Models	Deviance	dAIC	AICw
M.0 SpeedAvg ∼(1 ∣ id)	-129728	2322.82	0
M.1 SpeedAvg ∼Length + (1 ∣ id)	-131824	229.57	0
M.2 SpeedAvg ∼Length + music education + (1 ∣ id)	-131818	237.23	0
M.3 SpeedAvg ∼Length * music education + (1 ∣ id)	-131797	260.14	0
M.4 SpeedAvg ∼Length * music education + condition + (1 ∣ id)	-132061	0.00	1
M.5 SpeedAvg ∼Length * music education * condition + (1 ∣ id)	-131992	80.78	0

TABLE 5.18: GLMM comparison for working memory and isochrony in Study 5

Models	Deviance	dAIC	AICw
M.0 SpeedAvg ∼(1 ∣ id)	-129728	2365.38	0
M.1 SpeedAvg ∼Length + (1 ∣ id)	-131824	272.13	0
M.2 SpeedAvg ∼Length + working memory + (1 ∣ id)	-131823	275.23	0
M.3 SpeedAvg ∼Length * working memory + (1 ∣ id)	-131837	262.44	0
M.4 SpeedAvg ∼Length * working memory + condition + (1 ∣ id)	-132104	0.00	1
M.5 SpeedAvg ∼Length * working memory * condition + (1 ∣ id)	-132038	78.24	0

TABLE 5.19: GLMM comparison for isochrony in Study 5

Models	Deviance	dAIC	AICw
M.1 SpeedAvg ∼Length * dyslexia risk + condition + (1 | id)	-132078	25.52	0
M.2 SpeedAvg ∼Length * music education + condition + (1 | id)	-132061	42.56	0
M.3 SpeedAvg ∼Length * working memory + condition + (1 | id)	-132104	0	1

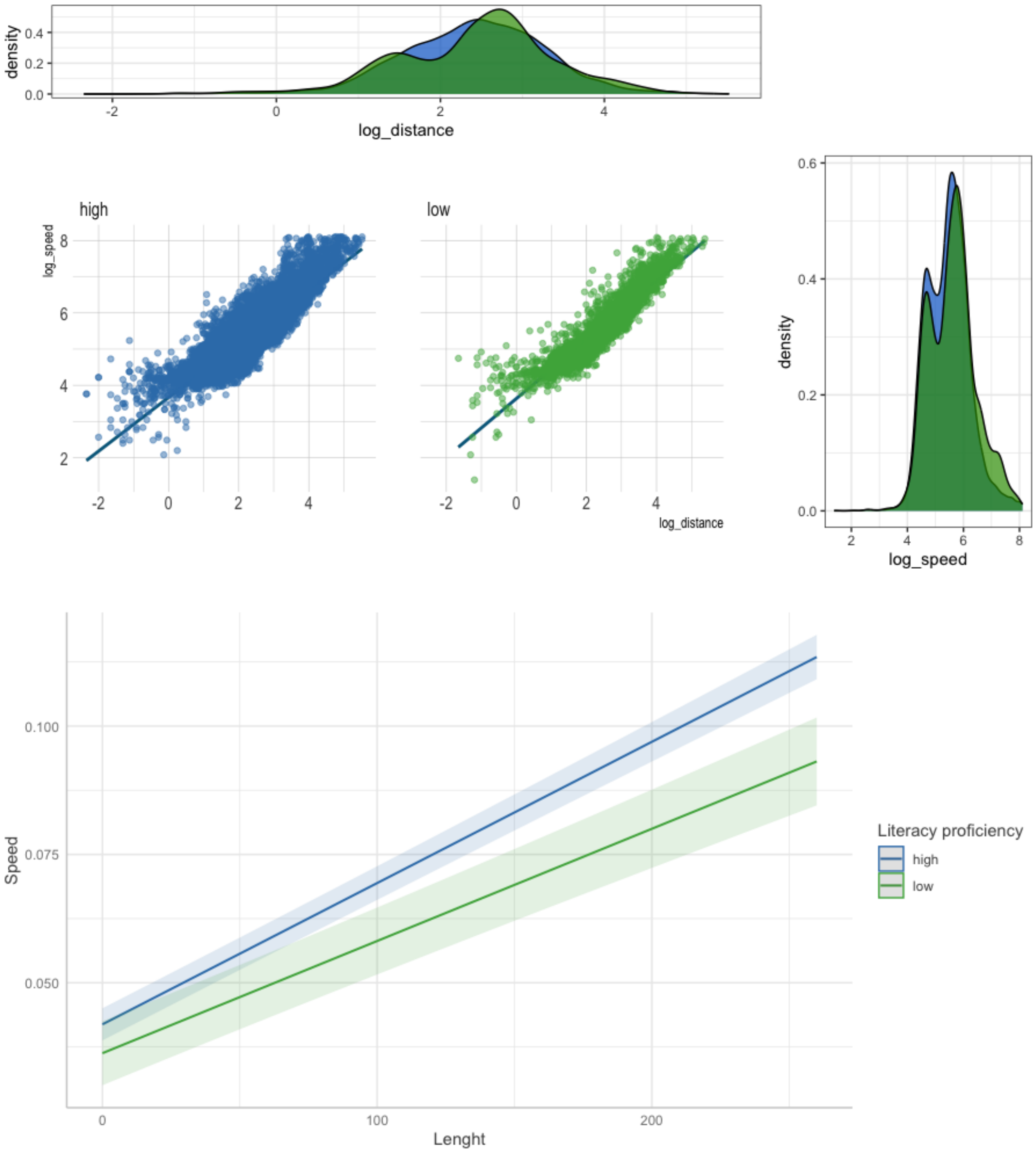


FIGURE 5.15: Descriptive statistics as log-log plot (natural logarithm) for low (blue) and high (green) levels of literacy proficiency (i.e., high or low) with marginal density plots (top). Effect plot of M4 predicting the effect of literacy proficiency on handwriting isochrony (i.e., ratio between the speed and the length of a trajectory; bottom).

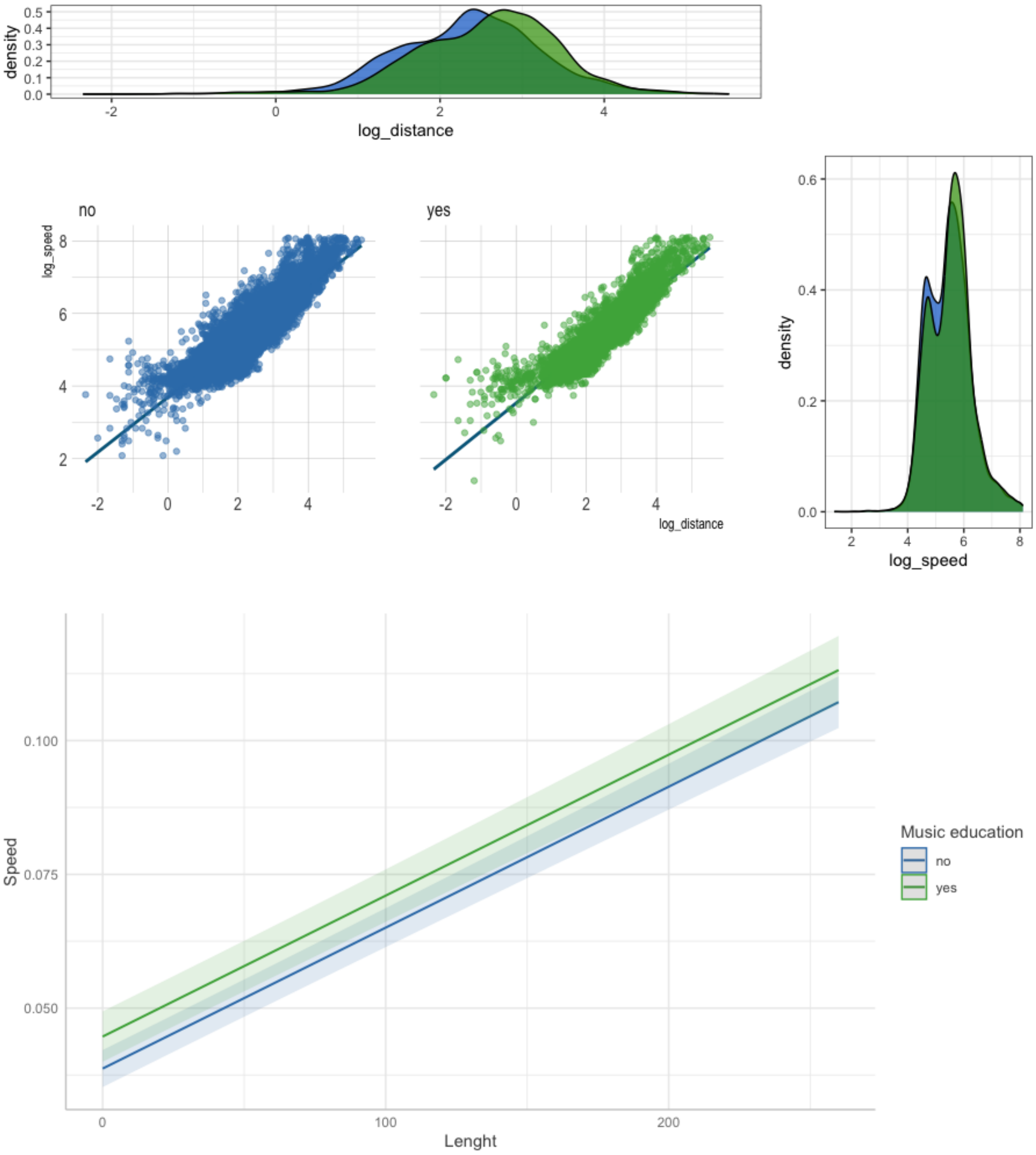


FIGURE 5.16: Descriptive statistics as log-log plot (natural logarithm) for low (blue) and high (green) level of music education (i.e., not received or received for more than 1 year) with marginal density plots (top). Effect plot of M4 predicting the effect of music education on handwriting isochrony (i.e., ratio between the speed and the length of a trajectory; bottom).

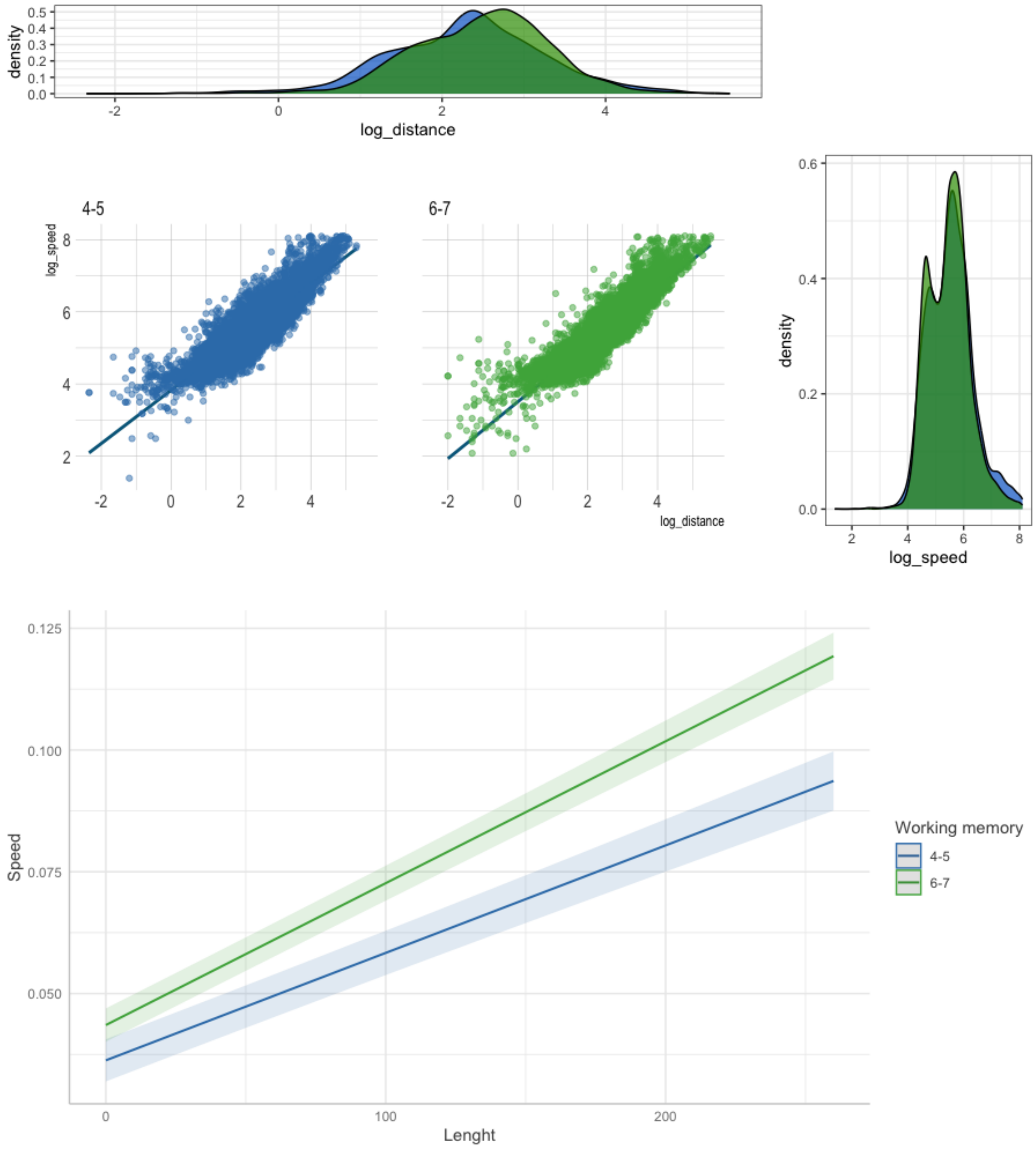


FIGURE 5.17: Descriptive statistics as log-log plot (natural logarithm) for low (bluee) and high (green) level of working memory (i.e., 4-5 or 6-7 score bin in the digit span test) with marginal density plots (top). Effect plot of M4 predicting the effect of working memory on handwriting isochrony (i.e., ratio between the speed and the length of a trajectory; bottom).

Interim discussion

The effect of rhythm in sustaining cognitive and phonological underpinnings of writing abilities has been investigated in this Study. Specifically, participants were asked to perform a dictation task i) in a baseline condition, ii) under articulatory suppression, and iii) cued by a regular rhythm. Performances were assessed in terms of product quality and quantitative measures including kinematic indexes (i.e., average speed and pressure of stroke) and adherence to isochrony (i.e., ratio between speed and length of stroke trajectory). A decrease in overall writing performance in the suppression condition was expected due to the interference of verbal articulation on the phonological loop. By contrast, the rhythmic condition was expected to contain the effect of the articulatory interference by cueing the dynamic attending of resources allocation via sensorimotor coupling, thus improving writing. Furthermore, low reading levels (assessed via a self-report screening for dyslexia in the first experimental session) were expected to predict lower writing performance. Specifically, poor readers were expected to fail to adhere to the isochrony principle of handwriting, replicating findings from Pagliarini et al. (2015). That is, participants with reading difficulties were expected to fail in enhancing the speed of movements for increasing length of trajectories.

Therefore, reduced performance in isochrony and kinematic indexes resulting in lower product quality were expected in comparison to proficient readers. This would confirm the hypothesis suggesting a common underpinning to the high comorbidity and similarities between dyslexic and dysgraphic profiles observed in atypical language development. Crucially, an external rhythmic cue was expected to sustain poor readers by acting on possible underlying impairments in timing processing hypothesized to be foundational in atypical language processing (Goswami, 2011; Ladányi et al., 2020; Lense et al., 2021). Consistently, long-term exposure to rhythm (assessed in terms of music education) was expected to predict higher writing performance across measures. Specifically, the effect of short and long-term exposure to rhythm was expected to improve writing performance by aiding the efficiency of dynamic resource allocation thus improving working memory and phonological processing.

Consistently, higher levels of working memory scores (assessed through digit span test in the first experimental session) were expected to predict better writing performance across measures. Therefore, this Study aimed at contributing to the literature identifying possible underlying difficulties in literacy development as well as at extending previous findings by exploring possible mechanisms whereby rhythm might improve writing.

Results on sixty young adults showed that reading abilities were the best predictor of product quality thus confirming the close link between dyslexia and dysgraphia observed in the literature (Döhla and Heim, 2016; Pagliarini et al., 2015). Furthermore, a strong involvement of rhythmic abilities underlying literacy difficulties is suggested by the results of isochrony analysis showing that poor readers fail to adhere with the rhythmic principles of handwriting, not improving movement speed when increasing in trajectory. Accordingly, poor reading abilities significantly predicted slower speed levels, as shown by the kinematic analysis. Therefore, this first result confirms predictions thus aligning with the literature suggesting a general timing deficit underlying poor literacy performance across language disorders (Goswami, 2011; Ladányi et al., 2020; Lense et al., 2021). Crucially, the regular rhythmic cue originally applied in this Study was found to significantly improve writing performances in poor readers. Specifically, the rhythmic cue significantly predicted a consistent improvement in product quality aligning performances of poor readers to those of proficient participants. Furthermore, whereas low reading levels significantly predicted a decrease in the average speed in the suppression condition, this was not the case for the rhythmic condition where a comparably increased average speed was observed. Therefore, if the present result confirms the existence of a core rhythmic deficit in handwriting for poor readers (i.e., a failure in complying with handwriting isochrony), it also suggests that an external rhythm might efficiently target this underlying impairment thus improving writing.

Consistently, a long-term exposure to rhythm significantly predicted an overall improvement in writing performance. Specifically, receiving a music education (i.e., coded as years of individual lessons with a music instrument) significantly predicted higher quality scores, especially when performing under articulatory suppression thus probably acting on the phonological and processing mechanisms underlying writing. Consistently, receiving a music education significantly predicted an increase in the velocity at which the isochrony principle was instantiated. Therefore, the present Study further confirms the core involvement of rhythm in literacy abilities, bringing new evidence regarding the effect of short- and long-term exposure to rhythm in the development of writing skills. Furthermore, the evidence of a rhythmic effect on the underlying difficulties in handwriting suggests that the current models framing the role of rhythm in language development might be extended to writing abilities including dysgraphic disorders. Accordingly, the prevalence of writing disorders is estimated to be around 7–15% among school-aged children, thus only slightly lowering the incidence of developmental dyslexia, which is about 17% (Döhla and Heim, 2016; Hawke et al., 2009; Shaywitz and Shaywitz, 2005). Moreover, the aforementioned rates of comorbidity result in an overall prevalence of learning disorders (including reading, writing, and counting difficulties) of about 5–15% worldwide (Association et al., 2020). Nevertheless, the cognitive profile of individuals with dysgraphia received significantly less attention compared to other atypical profiles (Döhla and Heim, 2016; Döhla et al., 2018; McCloskey and Rapp, 2017).

By contrast, in the present Study the interplay between phonological and cognitive mechanisms underlying writing difficulties has been investigated by exploring the predictive effect of working memory levels on handwriting. Results showed that higher levels of working memory (assessed through the digit span test in the first experimental session) significantly predicted an improvement in product quality and average speed of handwriting across conditions. Furthermore, whereas reading levels did not resulted as a significant predictor of average pen pressure in handwriting (Kushki et al., 2011; Pagliarini et al., 2015), working memory levels were the unique fixed effect resulting as a significant predictor of pen pressure in the kinematic analysis.

Specifically, lower levels of average pressure were significantly predicted by higher levels of working memory. Referring to the literature, the pressure exerted with the pen on the surface is primarily a sign of fatigue in handwriting (Mekyska et al., 2016; Asselborn et al., 2018). Therefore, the combination of increased speed and reduced pressure predicted in the kinematic analysis for higher levels of working memory can be interpreted as a facilitating effect in organizing the motor pattern of handwriting (Asselborn et al., 2018; Mekyska et al., 2016). Interestingly, increasing levels of average pressure were predicted in the rhythmic condition for lower working memory scores.

However, this trend resulted in an overall increase in product quality scores for lower working memory levels in the rhythmic condition. Therefore, this result might be interpreted as a possible compensatory mechanism while performing handwriting in the presence of low working memory resources. Consistently with this interpretation, a lack of automatism at the cognitive level, linked to possible impairments of procedural learning circuits in the cerebellum, has been framed in the literature as a possible mechanism in learning disorders (Nicolson et al., 2001; Nicolson and Fawcett, 2007). Specifically, individuals with dysgraphia are hypothesized to struggle in automatizing basic processes in handwriting thus requiring higher levels of cognitive load and working memory to perform writing (Nicolson and Fawcett, 2011). Consequently, individuals with learning impairments are hypothesized to master 'conscious compensatory mechanisms' to develop literacy skills (Silverman et al., 2008; Nicolson and Fawcett, 2007). Consistently with the present findings, individuals with ADHD and dysgraphia (Kroese et al., 2000; Mayes et al., 2000, two developmental disorders exhibiting high rates of comorbidity) were found to increase average pen pressure to gain control in handwriting, improving performances (Schoemaker et al., 2005; Adi-Japha et al., 2007; Van Gemmert and Van Galen, 1997). Furthermore, individuals with ADHD participating in the study from Adi-Japha et al. (2007) were typical readers, thus the difficulties in writing (and consequent compensatory mechanism) were not causally linked to a primary linguistic deficit, being rather explained in terms of attention and working memory load impacting language as a second factor (Adi-Japha et al., 2007). Similarly, the present findings suggest that working memory processes are involved in the general underpinnings of handwriting.

Moreover, previous findings are extended by bringing new evidence on the effect of a regular rhythm on this process. Specifically, this link holds on two findings. Firstly, presenting a regular rhythmic cue significantly sustained writing outcomes in individuals with low levels of working memory, probably through compensatory mechanisms in handwriting. Secondly, working memory scores resulted as the best predictor of isochrony in handwriting. Specifically, high levels of working memory significantly predicted a better adherence to the isochrony principle. Therefore, results from Study 5 suggest a positive effect of short- and long-term exposure to rhythm on the cognitive mechanisms underlying writing abilities including phonological processing, working memory, and motor control. Accordingly, the sensorimotor coupling elicited in the rhythmic condition is here hypothesized to sustain the dynamic allocation of attentional resources. Consequently, improvements in working memory and phonological processing result in a more efficient handwriting execution of internal motor scripts. Accordingly, the involvement of sensorimotor coupling as an underlying mechanism of rhythm processing is here suggested to be particularly effective in investigating and training writing abilities, given the intrinsic motor and linguistic nature of this literacy skill.

Conclusions to Study 5

In the present Study, the effect of sensorimotor rhythm on writing abilities has been investigated in young adults. Results support a significant effect of rhythm on the cognitive and motor abilities underlying handwriting. Based on this evidence, a sensorimotor engagement to rhythm is suggested to be a good candidate for investigating the effect of resources allocation and working memory on phonological processing and motor planning. Furthermore, the combined use of qualitative and quantitative measures of writing performance has revealed to be particularly informative about the underlying mechanisms supporting handwriting. For instance, beyond the classical legibility evaluation common across screening test for dysgraphia (Barnett et al., 2018; Tressoldi et al., 2013), kinematic and isochrony measures were collected by means of a digitalized iPad and a wireless Apple pen. Specifically, measures of average speed of stroke and average pen pressure (i.e., kinematic indexes) were informative about the effect of rhythm on phonological and working memory revealing an interaction with reading level and digit span scores.

Additionally, computing the ratio between average speed and length of trajectory was informative about the modulating effect of reading abilities, music education, and working memory on the adherence to the isochrony principle governing the rhythmic structure of handwriting. Specifically, reading abilities and working memory were positive predictors of the level of adherence, whereas music education was a positive predictor of the overall velocity whereby isochrony was instantiated. In conclusion, the present findings might inform the current models framing: i) atypical rhythm as a potential risk factor for language acquisition and ii) a functional link between general rhythmic abilities and language development by including new evidence on writing abilities and developmental dysgraphia. Furthermore, the present results might help in clarifying the effect of sensorimotor rhythm on cognitive and motor underpinning of handwriting thus advancing the current knowledge on typical and atypical writing development required for better targeting screening and supporting interventions.

5.3 Discussion and conclusion to Chapter five

Chapter five investigated the effect of sensorimotor rhythm on literacy skills. Specifically, sixty young adults participated in two experimental sessions investigating reading and writing abilities, by means of a lexical decision and a dictation task. The administration of both tasks was digitized to improve the qualitative and quantitative analysis of performance. Specifically, reading was evaluated in terms of accuracy and cognitive load (i.e., pupillometry) by means of an eye-tracking system; whereas writing abilities were evaluated in terms of product quality (i.e., legibility), kinematic indexes (i.e., average speed and average pressure of stroke), and adherence to isochrony (i.e., the ratio between speed and length of a movement) by means of an iPad and a wireless pen. Furthermore, difficulties in literacy abilities experienced in everyday life were investigated by means of a self-report screening for dyslexia. Additionally, previous exposure to rhythm was assessed in terms of years of music education. Lastly, participants were asked to perform a digit span test aimed at collecting information on their working memory level. Data were modelled by means of Generalized Mixed-Effects models (Nieuwenhuis et al., 2012; Fox, 2015) applied to estimate the effect of predictors while dealing for participant and item variance. The effect of different predictors was then explored following a hierarchical step-wise model comparison, while model comparison was performed using the likelihood ratio test (LTR), Akaike Information Criterion (Akaike, 1974, AIC), and AIC weight as indexes of the goodness of fit.

Overall, results support a positive effect of rhythm on reading and writing performance. Specifically, the effect of sensorimotor coupling to a regular rhythmic cue was found to significantly predict a decrease in the cognitive load required to poor readers when performing a demanding verbal task (i.e., lexical decision under articulatory suppression). Furthermore, it predicted a significant increase in product quality and average speed of handwriting, with comparable performances between poor and proficient readers in the rhythmic condition. Similarly, a positive effect of long-term engagement to a rhythmic activity (i.e., practicing with a music instrument) significantly predicted an increase in accuracy and a decrease in cognitive load in reading while entraining to a regular rhythmic cue.

Furthermore, receiving a music education predicted a significant increase in terms of product quality and velocity whereby isochrony was realized in handwriting thus supporting the link between long-term exposure to rhythm and literacy development. Specifically, the underlying mechanisms allowing for the transfer of benefit from rhythm to language are hypothesized to involve general cognitive mechanisms including working memory and dynamic resource allocation. Consistently, high levels of working memory predicted a significant improvement in accuracy and a decrease in cognitive load while reading as well as significantly better performance in terms of product quality, average speed and pressure, and adherence to isochrony in handwriting.

Overall, results support a positive effect of rhythm on reading and writing performance. Specifically, the effect of sensorimotor coupling to a regular rhythmic cue was found to significantly predict a decrease in the cognitive load required to poor readers when performing a demanding verbal task (i.e., lexical decision under articulatory suppression). Furthermore, it predicted a significant increase in product quality and average speed of handwriting, with comparable performances between poor and proficient readers in the rhythmic condition. Similarly, a positive effect of long-term engagement to a rhythmic activity (i.e., practicing with a music instrument) significantly predicted an increase in accuracy and a decrease in cognitive load in reading while entraining to a regular rhythmic cue. Furthermore, receiving a music education predicted a significant increase in terms of product quality and velocity whereby isochrony was realized in handwriting thus supporting the link between long-term exposure to rhythm and literacy development. Specifically, the underlying mechanisms allowing for the transfer of benefit from rhythm to language are hypothesized to involve general cognitive mechanisms including working memory and dynamic resource allocation. Consistently, high levels of working memory predicted a significant improvement in accuracy and a decrease in cognitive load while reading as well as significantly better performance in terms of product quality, average speed and pressure, and adherence to isochrony in handwriting.

Furthermore, working memory resulted as the best predictor of performance in the isochrony analysis thus revealing a core involvement in the rhythmic structure governing the motor planning of actions in writing. Accordingly, entrain movements to a rhythmic cue was found to be particularly effective in sustaining writing for low working memory levels, probably via kinematic compensatory mechanisms elicited to overcome an impaired automatizing process in motor planning. Therefore, results support the effect of short- and long-term sensorimotor engagement to rhythm on the underlying mechanisms supporting literacy abilities, with a particular efficacy observed in improving writing skills.

This result has been discussed considering the current models framing the effect of rhythm on language processing. Specifically, evidence on reading and writing difficulties were interpreted by referring to the TSF (Goswami, 2011) suggesting a processing deficit in the entrainment of neural oscillations underlying phonological processing in dyslexia. Accordingly, high dyslexia risk levels significantly predicted a strong interference of articulatory suppression on phonological processing in reading. Interestingly, difficulties in phonological processing were found to extend to writing production and this result was interpreted in light with the hypothesis of a general timing deficit underlying language-related impairments in development (Ladányi et al., 2020; Lense et al., 2021; Nayak et al., 2022). Furthermore, the increased effect of sensorimotor rhythm on writing is explained by the intrinsic coupling of motor and language abilities involved in handwriting. Consistently, this ability might benefit the most from a rhythmic exposure by activating the three mechanisms underlying rhythm processing across different domains, in line with the PRISM framework (Fiveash et al., 2021, i.e., fine auditory processing, neural entrainment, and sensorimotor coupling). Therefore, the present Chapter offers a composite evaluation of the effect of rhythm on cognitive and phonological mechanisms underlying literacy skills in young adults.

Interestingly, although the tested sample participating in Study 4 and 5 was identified as at typical development, at-risk scores in the self-report screening for dyslexia were observed and significantly contributed in explaining the performance across tasks. Similarly, variance in working memory and music education within the sample contributed to estimate the observed effects. Therefore, the instruments and methods applied in Study 4 and 5 were efficient in highlighting individual differences, letting the continuum of literacy abilities emerge. Furthermore, the effect of different contributing factors (i.e., task, stimuli, condition, reading level, working memory score, music education) was evaluated by means of mixed-effect models and model comparison. Specifically, this statistical approach allowed for testing the fixed effect of predictors while dealing with inter-individual and inter-item variance. Furthermore, digitizing traditional procedures for assessing reading and writing skills allowed for collecting a composite set of behavioural, physiological, kinematic, and self-report data useful for maximizing the analysis of the different aspects contributing to the phenomenon of interest.

This multi-method approach is particularly appropriate for studying the complex interplay dynamics occurring between basic mechanisms underlying high-level cognitive skills including language. Accordingly, perceptual, phonological, memory, attentive, and kinematic abilities were evaluated through Study 4 and 5 to investigate the mechanisms whereby rhythm might affect literacy. Therefore, this approach is suggested to significantly contribute to investigating the link between rhythm and language in typical and at-risk development. Consistently, bringing new findings to update the current models of language development is essential to inform clinicians and health professionals about risk factors and strengths to target within training or treatments. In conclusion, the effect of rhythm on high linguistic skills was investigated in this Chapter. In the next and last Chapter of this dissertation, the contributions of these findings (Study 4 and 5), together with the evidence on the link between VTS rhythm and language acquisition in infancy (Study 1, 2, and 3), will be discussed with the aim to trace a possible trajectory of the effect of body rhythm on language development.

Part III

DISCUSSION AND CONCLUSION

Chapter 6

Conclusion and Future Work

6.1 Summary of results

The short- and long-term influence of rhythm on the underlying mechanisms supporting reading and writing abilities was investigated in Chapter 5. Specifically, the observed improvements in literacy skills were discussed in terms of elicited sensorimotor entrainment of neural oscillations at Delta-band frequencies supporting phonological processing in reading (Study 4) and kinematic dynamics in handwriting (Study 5). Within the theoretical framework of TSF (Goswami, 2011), the present findings thus highlight the role of rhythm in improving the perceptual and cognitive mechanisms underlying reading abilities. Furthermore, Chapter 4 extends the TSF claims on timing deficits in dyslexia by bringing new evidence on the underlying mechanisms impairing timing organization of handwriting in dysgraphia. Accordingly, dysgraphia is an impacting learning disorder with an estimated prevalence of about 7–15% among school-aged children, known to share several commonalities and high comorbidity levels with dyslexia (Döhla and Heim, 2016; Landerl and Moll, 2010; Pagliarini et al., 2015). However, the dysgraphic profile has gained so far significantly less attention in research (Döhla et al., 2018), often being neglected by current models on rhythm and language development (Ladányi et al., 2021; Lense et al., 2021; Nayak et al., 2022; Goswami, 2011).

By contrast, Chapter 5 suggests that i) rhythm might play a crucial role in developing typical writing abilities, and ii) eliciting long- and short-term sensorimotor coupling to rhythm might be particularly effective in supporting atypical handwriting. Accordingly, findings from Study 4 and Study 5 contribute new evidence to the recently formulated frameworks on rhythm and language (Fiveash et al., 2021, e.g., the PRISM framework) by highlighting the role of sensorimotor coupling in supporting language development. Furthermore, it is worth noting the considerable variance in the reading and writing abilities assessed among the sixty young adults participating in Study 4 and 5. Interestingly, the investigated effect of rhythmic processing, phonological mechanisms, and working memory as well as the self-reported measures of reading difficulties and music education significantly contributed in explaining part of this variance.

Therefore, Chapter 4 might also contribute in reinforcing the call, made by the recently formulated models of atypical cognitive development (Ladányi et al., 2020; Lense et al., 2021; Nayak et al., 2022), to investigate atypical rhythm processing across populations as a potential risk factor for language-related impairments. Accordingly, clarifying the mechanisms whereby rhythm might affect language in typical, atypical, and at-risk populations is crucial in understanding and treating language disorder. Lastly, Chapter 5 clearly showed how i) the acquisition of literacy abilities – a milestone of language development – substantially holds on foundational, basic mechanisms including phonological processing, working memory, and attentional deployment; and ii) that rhythm is involved in all of these. Therefore, understanding the developmental pathway whereby rhythm might shape language through basic cognitive abilities is crucial to prevent impaired outcomes. Accordingly, the study of the effects of rhythm on language development reported in this thesis began with infancy.

Specifically, the ability to process VTS rhythms (Study 1), its link with early language processing (Study 2), and its possible application in supporting language development in infancy (Study 3) was the focus of Chapter 4. Overall, Chapter 4 aimed at investigating the role that rhythmic processing under the understudied sensory modality of vestibular, tactile, and sensorimotor sensing (Provasi et al., 2014; Bremner and Spence, 2017, VTS) might play in early language acquisition. Specifically, this Chapter is based on an extensive literature review suggesting a primary role of early VTS experiences in shaping the perceptual and cognitive processing of rhythm (see Chapter 2). Specifically, infants were hypothesized to be primed by the prenatal exposure to the rhythms – conveyed via vibrations to the body and bone-conduction to the inner ear – produced by sounds and movements. Therefore, infants were expected to be highly familiar with rhythmic signals presented through the VTS modality. Consequently, the Infant vibrotactile system was originally designed to test this hypothesis presenting infants with sounds translated into vibrations. Crucially, this device was based on the current advances in the fields of haptic technology and computational sonology to maximize the signal features crucial in sound processing. That is, different frequency bands were presented to different portions of the back of the participants in order to maintain the topographic organization in band frequencies operated by the human cochlea (Karam et al., 2009b).

Therefore, the resulting signal was characterized by amplitude and frequency information continuously provided to the somatosensory system via mechanoreceptors, thus resulting in an immersive experience. The development of the Infant vibrotactile system was motivated by the aim to enrich the field of infant research with the technological advances from the haptic and music technology fields as well as to provide a relevant application for these advances thus informing the technology field in return. In fact, the Infant vibrotactile system was to our knowledge the first vibrotactile system with a multi-channel filtering of band frequencies specifically suited for the infant population. Furthermore, the Infant vibrotactile system was integrated with an eye-tracker to present infants with graze-triggered paradigms while continuously collecting data on variations in pupil diameter as an index of cognitive processing and attentional deployment (Mathôt, 2018; Hepach and Westermann, 2016).

By doing so, it was possible to i) test rhythmic abilities under the VTS modality in verbal as in preverbal infants, and ii) focus on the attentional processes (i.e., cognitive pupillometry) underlying the behaviors (i.e., looking times) displayed toward stimuli. The last methodological improvement applied in Chapter 4 was to significantly extend the age range of participants, with respect to previous studies, in order to trace the developmental changes occurring through infancy, from 7 to 30 months. Accordingly, the effect of predictors (i.e., trial type and condition) on dependent variables (i.e., pupil variation and proportions) was modeled by means of generalized mixed-effect models on the continuous course of time (GAMM) and age (GLMM). Therefore, by means of this experimental setting and design, and the statistical approach, infant VTS rhythmic abilities and their role in early language processing were investigated revealing interesting developmental changes through the age range.

Specifically, Study 1 showed that infants are able to discriminate between VTS rhythms based on the underlying metrical structure. Furthermore, the VTS rhythmic abilities assessed in Study 1 were found to significantly predict patterns of phonological and prosodic processing in Study 2. Lastly, Study 3 brought preliminary findings on cross-sensory improvements in processing auditory presented linguistic stimuli matching the meter of a VTS rhythmic prime. Furthermore, developmental changes through age showed that younger infants allocated increased attention toward familiar stimuli whereas older infants allocated increased attention toward novel ones. This finding has been interpreted as a sign of maturational changes in information processing: that is, younger infants might need more time to form a stable representation of the encoded information (i.e., mainly directing attention toward familiar stimuli), whereas older infants might be faster thus allowing the attentional system to switch the focus of attention toward new stimuli (DePaolis et al., 2016; Hunter and Ames, 1988). Interestingly, this developmental pattern characterizes the deployment of attention toward both VTS rhythm and auditory linguistic stimuli (Study 1 and 2). Therefore, the developmental stage of resource allocation in one domain (e.g., rhythm) or modality (e.g., VTS) might be informative about the other (i.e., language; auditory).

Accordingly, proportional increases in pupil diameter toward novel stimuli in the VTS rhythmic task significantly predicted proportional increases in pupil diameter toward novel stimuli in the linguistic task (Study 2) and it was found to increase with age. Moreover, phonological features in linguistic stimuli were found to gain increasing attention with age, whereas prosodic features were found to drive more attention earlier in age, paralleling the findings from the developmental stages of language acquisition (Kuhl, 2004; Gervain, 2018; Jusczyk et al., 1999; Seidl and Cristià, 2008). Interestingly, a domain related pattern of development was found in rhythm processing as well. Specifically, familiarizing with a double meter significantly predicted increasing discrimination abilities with age compared to a triple meter. This facilitating effect of double meter has been explained in terms of i) a processing bias toward binary rhythmic patterns (Ullal-Gupta et al., 2013; Teie, 2016, governing biological function as breathing rate, heart beat, or walking) and ii) an enculturation process toward the rhythmic pattern predominant in the native environment (Hannon and Johnson, 2005; Hannon and Trainor, 2007; Kalender et al., 2013, i.e., double meter in Western music).

Therefore, Chapter 4 and Chapter 5 together brought innovative evidence on the bodily-rooted influences of rhythm on language processing along development. Specifically, the effect of the prenatal experiences with VTS rhythms was observed in shaping the encoding abilities for rhythmic and linguistic stimuli through infancy. Specifically, from the broad-based processing of meters and suprasegmental features, to the specialized phonological and culture-specific tuning, infants have been found to go through maturational (i.e., from perceptual bias to information processing) and environmental (i.e., from broad-based to culture-specific) changes in the development of rhythm and language.

In Chapter 5, the impact of these basic abilities on the development of literacy skills was assessed through reading and writing tests. Going from a passive (i.e., VTS rhythmic perception) to an active (i.e., sensorimotor rhythmic entrainment) involvement of the body, rhythmic skills were found to play a role in entraining and handwriting dynamics supporting reading and writing abilities in young adults. Consistently, the effect of experience passive (Phillips-Silver and Trainor, 2005; Rocha et al., 2021) and active (Phillips-Silver and Trainor, 2007) movement was found to elicit VTS sensing thus influencing rhythm encoding in infants and adults (Phillips-Silver and Trainor, 2008; Tichko et al., 2021, 2022; Trainor et al., 2009). Crucially, the present dissertation represents the first attempt to our knowledge to trace a developmental pathway of the embodied influences of rhythm on language development, from the prenatal experience with VTS rhythms to the high-level sensorimotor influences on literacy skills.

6.2 Limits and future directions

Importantly, the presented results are the outcome of a journey that constantly required improvements and changes. Still, some limitations might arise. Therefore, further research is needed to replicate, corroborate, and extend the present findings trying to answer the left open questions. For instance, testing infants from non-Western countries will be useful to disentangle whether results are determined by music enculturation and language-specific processes or are rather perceptual universals. Furthermore, through an ongoing collaboration with the Speech Acquisition Lab at the Universitat Pompeu Fabra (Barcelona, Spain), I had the opportunity to explore the influence and extent of bilingual backgrounds on language acquisition. Therefore, investigating body rhythm in bilingual populations might be informative since they are exposed to the same source of biological rhythms but usually grow up in mixed music, cultural, and linguistic environments. Moreover, the limitations due to the COVID-19 prevented the implementation of a longitudinal design. However, follow-up infants would be useful to investigate the predictive strength of early rhythmic abilities on subsequent steps of language development. Nevertheless, the impossibility to pursue the lab-based research prompted us to bring the research outside laboratories and meet infants in familiar environments.

Thanks to the collaboration with the Cooperative Now Project, it was indeed possible to test infants across kindergartens in Padua and next municipalities, thus preventing the project from stopping. Furthermore, the forced pause to data collection during the first months of lockdown let the extensive literature review produce significant improvements from theoretical and technological sides. For instance, with the BabyLab of the University of Padua (IT) it was possible to implement an online eye-tracking data collection on the Labvanced platform. Similarly, parent-report and self-report questionnaires (e.g., the MacArthur–Bates Communicative Development Inventory, the Music@Home questionnaire, the Revised Checklist for Dyslexia) were digitized for online administration on the Qualtrics platform.

Therefore, future studies will apply a larger set of background measures to better estimate the contribution of environmental and daily-life variables. Another strict limitation to this project due to COVID-19 was the postponing of the originally planned data collection with deaf infants. In fact, testing the possible applications of the Infant vibrotactile system to support rhythm and language processing in deaf infants was one of the main aims behind this project.

6.3 Conclusions

To conclude, a final remark on the core involvement of rhythm and sensorimotor processes in language acquisition is worth doing. Specifically, further research is needed to bring evidence-based knowledge to clinicians and health professionals to include rhythmic-based screening and training for language and cognitive development. Additionally, education systems and schools will benefit from further research on the effectiveness of rhythm to promote development, not only to inform professionals and caregivers about early risk factors and good practice; rather, decision makers and directors need evidence-based, high-quality findings to apply for the financial support required to achieve significant changes in the education system. For instance, evidence on the impact of active practice with a musical instrument on language and cognitive performance is growing; however, the access to music education is still not guaranteed to everyone. By contrast, the art system provides every year well-trained musicians who would significantly enrich the formative offer of the public school system. Similarly, developing high-quality practices based on music, rhythm, and motor activity will offer culture-free and language-free ways to connect with neurodivergent profiles thus helping to design individualized and collective interventions for answering the specific needs of the community. Therefore, this work aimed to be a first, tiny step into this collective challenge.

Bibliography

Abboub, N., Nazzi, T., and Gervain, J. (2016). Prosodic grouping at birth. *Brain and language*, 162:46–59.

Abu-Zhaya, R., Seidl, A., and Cristia, A. (2017). Multimodal infant-directed communication: How caregivers combine tactile and linguistic cues. *Journal of Child Language*, 44(5):1088–1116.

Accardo, A. P., Genna, M., and Borean, M. (2013). Development, maturation and learning influence on handwriting kinematics. *Human movement science*, 32(1):136–146.

Adams, I. L., Lust, J. M., Wilson, P. H., and Steenbergen, B. (2017). Testing predictive control of movement in children with developmental coordination disorder using converging operations. *British Journal of Psychology*, 108(1):73–90.

Adi-Japha, E., Landau, Y. E., Frenkel, L., Teicher, M., Gross-Tsur, V., and Shalev, R. S. (2007). Adhd and dysgraphia: underlying mechanisms. *Cortex*, 43(6):700–709.

Akaike, H. (1974). A new look at the statistical model identification. *IEEE transactions on automatic control*, 19(6):716–723.

Álamos Gómez, J. E. and Tejada, J. (2020). Interrelaciones entre acción y cognición. aportaciones de la neurociencia a la educación rítmico-musical. *Opus, 2020, vol. 26, num. 2, p. 1-21.*

Allen, S. and Casey, J. (2017). Developmental coordination disorders and sensory processing and integration: Incidence, associations and comorbidities. *British journal of occupational therapy*, 80(9):549–557.

Alm, P. A. (2004). Stuttering and the basal ganglia circuits: a critical review of possible relations. *Journal of communication disorders*, 37(4):325–369.

Ammirante, P., Patel, A. D., and Russo, F. A. (2016). Synchronizing to auditory and tactile metronomes: a test of the auditory-motor enhancement hypothesis. *Psychonomic bulletin & review*, 23(6):1882–1890.

Ammirante, P., Russo, F. A., Good, A., and Fels, D. I. (2013). Feeling voices. *PLoS One*, 8(1):e53585.

Andreola, C., Mascheretti, S., Belotti, R., Ogliari, A., Marino, C., Battaglia, M., and Scaini, S. (2021). The heritability of reading and reading-related neurocognitive components: A multi-level meta-analysis. *Neuroscience & Biobehavioral Reviews*, 121:175–200.

Anvari, S. H., Trainor, L. J., Woodside, J., and Levy, B. A. (2002). Relations among musical skills, phonological processing, and early reading ability in preschool children. *Journal of experimental child psychology*, 83(2):111–130.

Arias-Trejo, N., Angulo-Chavira, A. Q., Avila-Varela, D. S., Chua-Rodriguez, F., and Mani, N. (2022). Developmental changes in phonological and semantic priming effects in spanish-speaking toddlers. *Developmental Psychology*, 58(2):236.

Arnal, L. H. and Giraud, A.-L. (2012). Cortical oscillations and sensory predictions. *Trends in cognitive sciences*, 16(7):390–398.

Arriaga, G., Zhou, E. P., and Jarvis, E. D. (2012). Of mice, birds, and men: the mouse ultrasonic song system has some features similar to humans and song-learning birds.

Arvaniti, A. (2009). Rhythm, timing and the timing of rhythm. *Phonetica*, 66(1-2):46–63.

Arvaniti, A. (2012). The usefulness of metrics in the quantification of speech rhythm. *Journal of Phonetics*, 40(3):351–373.

Assaneo, M. F. and Poeppel, D. (2018). The coupling between auditory and motor cortices is rate-restricted: Evidence for an intrinsic speech-motor rhythm. *Science advances*, 4(2):eaao3842.

Assaneo, M. F., Ripollés, P., Orpella, J., Lin, W. M., de Diego-Balaguer, R., and Poeppel, D. (2019). Spontaneous synchronization to speech reveals neural mechanisms facilitating language learning. *Nature neuroscience*, 22(4):627–632.

Asselborn, T., Guneysu, A., Mrini, K., Yadollahi, E., Ozgur, A., Johal, W., and Dillenbourg, P. (2018). Bringing letters to life: handwriting with haptic-enabled tangible robots. In *Proceedings of the 17th ACM conference on interaction design and children*, pages 219–230.

Association, A. P. et al. (2020). *The American Psychiatric Association practice guideline for the treatment of patients with schizophrenia*. American Psychiatric Pub.

Astheimer, L. B. and Sanders, L. D. (2009). Listeners modulate temporally selective attention during natural speech processing. *Biological psychology*, 80(1):23–34.

Aubanel, V., Davis, C., and Kim, J. (2016). Exploring the role of brain oscillations in speech perception in noise: intelligibility of isochronously retimed speech. *Frontiers in human neuroscience*, 10:430.

Baayen, H., Vasishth, S., Kliegl, R., and Bates, D. (2017). The cave of shadows: Addressing the human factor with generalized additive mixed models. *Journal of Memory and Language*, 94:206–234.

Baddeley, A. (1993). Short-term phonological memory and long-term learning: A single case study. *European Journal of Cognitive Psychology*, 5(2):129–148.

Baddeley, A., Lewis, V., and Vallar, G. (1984). Exploring the articulatory loop. *The Quarterly Journal of Experimental Psychology Section A*, 36(2):233–252.

Barnes, R. and Jones, M. R. (2000). Expectancy, attention, and time. *Cognitive psychology*, 41(3):254–311.

Barnett, A. L., Prunty, M., and Rosenblum, S. (2018). Development of the handwriting legibility scale (hls): A preliminary examination of reliability and validity. *Research in developmental disabilities*, 72:240–247.

Bates, D., Mächler, M., Bolker, B., and Walker, S. (2014). Fitting linear mixed-effects models using lme4. *arXiv preprint arXiv:1406.5823*.

Beatty, J., Lucero-Wagoner, B., et al. (2000). The pupillary system. *Handbook of psychophysiology*, 2(142-162).

Bedoin, N., Brisseau, L., Molinier, P., Roch, D., and Tillmann, B. (2016). Temporally regular musical primes facilitate subsequent syntax processing in children with specific language impairment. *Frontiers in Neuroscience*, 10:245.

Beker, S., Foxe, J. J., and Molholm, S. (2021). Oscillatory entrainment mechanisms and anticipatory predictive processes in children with autism spectrum disorder. *Journal of neurophysiology*, 126(5):1783–1798.

Benasich, A. A. and Tallal, P. (2002). Infant discrimination of rapid auditory cues predicts later language impairment. *Behavioural brain research*, 136(1):31–49.

Bergeson, T. R. and Trehub, S. E. (2006). Infants perception of rhythmic patterns. *Music Perception*, 23(4):345–360.

Bernard, C. and Gervain, J. (2012). Prosodic cues to word order: what level of representation? *Frontiers in Psychology*, 3:451.

Bernard, C., Monnoyer, J., Wiertlewski, M., and Ystad, S. (2022). Rhythm perception is shared between audio and haptics. *Scientific Reports*, 12(1):1–12.

Bertoncini, J., Floccia, C., Nazzi, T., and Mehler, J. (1995). Morae and syllables: Rhythmical basis of speech representations in neonates. *Language and speech*, 38(4):311–329.

Besson, M., Chobert, J., and Marie, C. (2011). Transfer of training between music and speech: common processing, attention, and memory. *Frontiers in psychology*, 2:94.

Besson, M., Schön, D., Moreno, S., Santos, A., and Magne, C. (2007). Influence of musical expertise and musical training on pitch processing in music and language. *Restorative neurology and neuroscience*, 25(3-4):399–410.

Bhatia, A. et al. (2013). Binary musical bias in irreguilar meters.

Bishop, D. V. (2017). Why is it so hard to reach agreement on terminology? the case of developmental language disorder (dld). *International journal of language & communication disorders*, 52(6):671–680.

Bispham, J. (2006). Rhythm in music: What is it? who has it? and why? *Music perception*, 24(2):125–134.

Blood, A. J. and Zatorre, R. J. (2001). Intensely pleasurable responses to music correlate with activity in brain regions implicated in reward and emotion. *Proceedings of the national academy of sciences*, 98(20):11818–11823.

Bolger, D., Trost, W., and Schön, D. (2013). Rhythm implicitly affects temporal orienting of attention across modalities. *Acta psychologica*, 142(2):238–244.

Bolton, T. L. (1894). Rhythm. *The american journal of psychology*, 6(2):145–238.

Bonacina, S., Krizman, J., White-Schwoch, T., Nicol, T., and Kraus, N. (2020). Distinct rhythmic abilities align with phonological awareness and rapid naming in school-age children. *Cognitive Processing*, 21(4):575–581.

Bouloukou, F., Marin-Diaz, V., and Jimenez-Fanjul, N. (2021). Effects of an interventional music program on learning skills of primary-school students with dyslexia. *International Journal of Education and Practice*, 9(3):456–467.

Brandon, M. and Saffran, J. R. (2011). Apparent motion enhances visual rhythm discrimination in infancy. *Attention, Perception, & Psychophysics*, 73(4):1016–1020.

Brandt, A., Gebrian, M., and Slevc, L. R. (2012). Music and early language acquisition. *Frontiers in psychology*, 3:327.

Branje, C., Nespoil, G., Russo, F., and Fels, D. I. (2014). The effect of vibrotactile stimulation on the emotional response to horror films. *Computers in Entertainment (CIE)*, 11(1):1–13.

Bremner, A. and Spence, C. (2017). The development of tactile perception. *Advances in child development and behavior*, 52:227–268.

Brochard, R., Tassin, M., and Zagar, D. (2013). Got rhythm… for better and for worse. cross-modal effects of auditory rhythm on visual word recognition. *Cognition*, 127(2):214–219.

Brown, L. L. (2012). The benefits of music education. *PBS KIDS for Parents*.

Bruderer, A. G., Danielson, D. K., Kandhadai, P., and Werker, J. F. (2015). Sensorimotor influences on speech perception in infancy. *Proceedings of the National Academy of Sciences*, 112(44):13531–13536.

Buck, J. B. (1938). Synchronous rhythmic flashing of fireflies. *The Quarterly Review of Biology*, 13(3):301–314.

Burani, C., Paizi, D., and Sulpizio, S. (2014). Stress assignment in reading italian: Friendship outweighs dominance. *Memory & Cognition*, 42(4):662–675.

Buzsaki, G. (2006). *Rhythms of the Brain*. Oxford university press.

Caçola, P. (2016). Physical and mental health of children with developmental coordination disorder. *Frontiers in public health*, 4:224.

Caetano, G. and Jousmäki, V. (2006). Evidence of vibrotactile input to human auditory cortex. *Neuroimage*, 29(1):15–28.

Calignano, G., Dispaldro, M., Russo, S., and Valenza, E. (2021a). Attentional engagement during syllable discrimination: The role of salient prosodic cues in 6-to 8-month-old infants. *Infant Behavior and Development*, 62:101504.

Calignano, G., Valenza, E., Vespignani, F., Russo, S., and Sulpizio, S. (2021b). The unique role of novel linguistic labels on the disengagement of visual attention. *Quarterly Journal of Experimental Psychology*, 74(10):1755–1772.

Cancer, A., Stievano, G., Pace, G., Colombo, A., and Antonietti, A. (2019). Cognitive processes underlying reading improvement during a rhythm-based intervention. a small-scale investigation of italian children with dyslexia. *Children*, 6(8):91.

Canette, L.-H., Bedoin, N., Lalitte, P., Bigand, E., and Tillmann, B. (2019). The regularity of rhythmic primes influences syntax processing in adults. *Auditory Perception & Cognition*, 2(3):163–179.

Canette, L.-H., Fiveash, A., Krzonowski, J., Corneyllie, A., Lalitte, P., Thompson, D., Trainor, L., Bedoin, N., and Tillmann, B. (2020). Regular rhythmic primes boost p600 in grammatical error processing in dyslexic adults and matched controls. *Neuropsychologia*, 138:107324.

Cannon, J. J. and Patel, A. D. (2021). How beat perception co-opts motor neurophysiology. *Trends in Cognitive Sciences*, 25(2):137–150.

Cantiani, C., Ortiz-Mantilla, S., Riva, V., Piazza, C., Bettoni, R., Musacchia, G., Molteni, M., Marino, C., and Benasich, A. A. (2019). Reduced left-lateralized pattern of event-related eeg oscillations in infants at familial risk for language and learning impairment. *Neuroimage: clinical*, 22:101778.

Carlson, G. N. (1989). On the semantic composition of english generic sentences. In *Properties, types and meaning*, pages 167–192. Springer.

Carrer, L. R. J. (2015). Music and sound in time processing of children with adhd. *Frontiers in psychiatry*, 6:127.

Cason, N., Astésano, C., and Schön, D. (2015a). Bridging music and speech rhythm: Rhythmic priming and audio–motor training affect speech perception. *Acta psychologica*, 155:43–50.

Cason, N., Hidalgo, C., Isoard, F., Roman, S., and Schön, D. (2015b). Rhythmic priming enhances speech production abilities: Evidence from prelingually deaf children. *Neuropsychology*, 29(1):102.

Cason, N. and Schön, D. (2012). Rhythmic priming enhances the phonological processing of speech. *Neuropsychologia*, 50(11):2652–2658.

Chang, A., Li, Y.-C., Chan, J. F., Dotov, D. G., Cairney, J., and Trainor, L. J. (2021). Inferior auditory time perception in children with motor difficulties. *Child Development*, 92(5):e907–e923.

Chang, S.-E., Chow, H. M., Wieland, E. A., and McAuley, J. D. (2016). Relation between functional connectivity and rhythm discrimination in children who do and do not stutter. *NeuroImage: Clinical*, 12:442–450.

Chen, J. L., Penhune, V. B., and Zatorre, R. J. (2008). Listening to musical rhythms recruits motor regions of the brain. *Cerebral cortex*, 18(12):2844–2854.

Cheng-Lai, A., Li-Tsang, C. W., Chan, A. H., and Lo, A. G. (2013). Writing to dictation and handwriting performance among chinese children with dyslexia: Relationships with orthographic knowledge and perceptual-motor skills. *Research in developmental disabilities*, 34(10):3372–3383.

Chern, A., Tillmann, B., Vaughan, C., and Gordon, R. L. (2018). New evidence of a rhythmic priming effect that enhances grammaticality judgments in children. *Journal of experimental child psychology*, 173:371–379.

Chobert, J., François, C., Velay, J.-L., and 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.

Choi, J.-Y., Kwon, E.-J., Lee, J.-Y., Song, J.-M., Kim, H.-J., and Kim, J.-S. (2021). Vestibular perception in time and space during whole-body rotation in humans. *The Cerebellum*, 20(4):509–517.

Christophe, T., Guasti, M., and Nespor, A. (1997). Reflections on phonological bootstrapping: Its role for lexical and syntactic acquisition. *Language and cognitive processes*, 12(5-6):585–612.

Cito, G., Luisi, S., Mezzesimi, A., Cavicchioli, C., Calonaci, G., and Petraglia, F. (2005). Maternal position during non-stress test and fetal heart rate patterns. *Acta obstetricia et gynecologica Scandinavica*, 84(4):335–338.

Cleveland, W. S. and Devlin, S. J. (1988). Locally weighted regression: an approach to regression analysis by local fitting. *Journal of the American statistical association*, 83(403):596–610.

Cogo-Moreira, H., de Avila, C. R. B., Ploubidis, G. B., and Mari, J. d. J. (2013). Effectiveness of music education for the improvement of reading skills and academic achievement in young poor readers: a pragmatic cluster-randomized, controlled clinical trial. *PloS one*, 8(3):e59984.

Colling, L. J., Noble, H. L., and Goswami, U. (2017). Neural entrainment and sensorimotor synchronization to the beat in children with developmental dyslexia: An eeg study. *Frontiers in Neuroscience*, 11:360.

Colombo, J. and Bundy, R. S. (1983). Infant response to auditory familiarity and novelty. *Infant Behavior and Development*, 6(2-3):305–311.

Colombo, L., Fudio, S., and Mosna, G. (2009). Phonological and working memory mechanisms involved in written spelling. *European Journal of Cognitive Psychology*, 21(6):837–861.

Coltheart, M., Besner, D., Jonasson, J. T., and Davelaar, E. (1979). Phonological encoding in the lexical decision task. *The Quarterly Journal of Experimental Psychology*, 31(3):489–507.

Conway, C. M., Pisoni, D. B., and Kronenberger, W. G. (2009). The importance of sound for cognitive sequencing abilities: The auditory scaffolding hypothesis. *Current directions in psychological science*, 18(5):275–279.

Cooper, R. P. and Aslin, R. N. (1990). Preference for infant-directed speech in the first month after birth. *Child development*, 61(5):1584–1595.

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

Coull, J. T. (2004). fmri studies of temporal attention: allocating attention within, or towards, time. *Cognitive Brain Research*, 21(2):216–226.

Craig, A. D. (2009). How do you feel—now? the anterior insula and human awareness. *Nature reviews neuroscience*, 10(1):59–70.

Crommett, L. E., Madala, D., and Yau, J. M. (2019). Multisensory perceptual interactions between higher-order temporal frequency signals. *Journal of Experimental Psychology: General*, 148(7):1124.

Crommett, L. E., Pérez-Bellido, A., and Yau, J. M. (2017). Auditory adaptation improves tactile frequency perception. *Journal of neurophysiology*, 117(3):1352–1362.

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

Curtin, S. and Archer, S. (2015). Speech perception. *The Cambridge handbook of child language*, pages 137–158.

Custode, S. A. and Tamis-LeMonda, C. (2020). Cracking the code: Social and contextual cues to language input in the home environment. *Infancy*, 25(6):809–826.

Cutini, S., Szűcs, D., Mead, N., Huss, M., and Goswami, U. (2016). Atypical right hemisphere response to slow temporal modulations in children with developmental dyslexia. *Neuroimage*, 143:40–49.

Cutler, A. (1984). Stress and accent in language production and understanding. *Intonation, accent, and rhythm: Studies in discourse phonology*, 8:77–90.

Cutler, A., Dahan, D., and Van Donselaar, W. (1997). Prosody in the comprehension of spoken language: A literature review. *Language and speech*, 40(2):141–201.

Dalla Bella, S., Farrugia, N., Benoit, C.-E., Begel, V., Verga, L., Harding, E., and Kotz, S. A. (2017). Baasta: Battery for the assessment of auditory sensorimotor and timing abilities. *Behavior Research Methods*, 49(3):1128–1145.

Dalmaijer, E. S., Mathôt, S., and Van der Stigchel, S. (2014). Pygaze: An open-source, cross-platform toolbox for minimal-effort programming of eyetracking experiments. *Behavior research methods*, 46(4):913–921.

Damasio, A. R. (1994). Descartes' error and the future of human life. *Scientific American*, 271(4):144–144.

Damasio, A. R. (2001). Descartes error revisited. *Journal of the History of the Neurosciences*, 10(2):192–194.

David, D., Wade-Woolley, L., Kirby, J. R., and Smithrim, K. (2007). Rhythm and reading development in school-age children: A longitudinal study. *Journal of Research in Reading*, 30(2):169–183.

de Castelnau, P., Albaret, J.-M., Chaix, Y., and Zanone, P.-G. (2007). Developmental coordination disorder pertains to a deficit in perceptuo-motor synchronization independent of attentional capacities. *Human Movement Science*, 26(3):477–490.

DeCasper, A. J. and Fifer, W. P. (1980). Of human bonding: Newborns prefer their mothers' voices. *Science*, 208(4448):1174–1176.

DeCasper, A. J., Lecanuet, J.-P., Busnel, M.-C., Granier-Deferre, C., and Maugeais, R. (1994). Fetal reactions to recurrent maternal speech. *Infant behavior and development*, 17(2):159–164.

DeCasper, A. J. and Sigafoos, A. D. (1983). The intrauterine heartbeat: A potent reinforcer for newborns. *Infant Behavior & Development*.

DeCasper, A. J. and Spence, M. J. (1986). Prenatal maternal speech influences newborns' perception of speech sounds. *Infant behavior and Development*, 9(2):133–150.

Dellatolas, G., Watier, L., Le Normand, M.-T., Lubart, T., and Chevrie-Muller, C. (2009). Rhythm reproduction in kindergarten, reading performance at second grade, and developmental dyslexia theories. *Archives of Clinical Neuropsychology*, 24(6):555–563.

Delle Luche, C., Durrant, S., Poltrock, S., and Floccia, C. (2015). A methodological investigation of the intermodal preferential looking paradigm: Methods of analyses, picture selection and data rejection criteria. *Infant Behavior and Development*, 40:151–172.

DePaolis, R. A., Keren-Portnoy, T., and Vihman, M. (2016). Making sense of infant familiarity and novelty responses to words at lexical onset. *Frontiers in psychology*, 7:715.

Di Brina, C., Niels, R., Overvelde, A., Levi, G., and Hulstijn, W. (2008). Dynamic time warping: A new method in the study of poor handwriting. *Human movement science*, 27(2):242–255.

d'Imperio, M. and Rosenthall, S. (1999). Phonetics and phonology of main stress in italian. *Phonology*, 16(1):1–28.

Ding, N., Patel, A. D., Chen, L., Butler, H., Luo, C., and Poeppel, D. (2017). Temporal modulations in speech and music. *Neuroscience & Biobehavioral Reviews*, 81:181–187.

Doheny, L., Hurwitz, S., Insoft, R., Ringer, S., and Lahav, A. (2012). Exposure to biological maternal sounds improves cardiorespiratory regulation in extremely preterm infants. *The journal of maternal-fetal & neonatal medicine*, 25(9):1591–1594.

Döhla, D. and Heim, S. (2016). Developmental dyslexia and dysgraphia: What can we learn from the one about the other? *Frontiers in psychology*, 6:2045.

Döhla, D., Willmes, K., and Heim, S. (2018). Cognitive profiles of developmental dysgraphia. *Frontiers in psychology*, 9:2006.

Doupe, A. J. and Kuhl, P. K. (1999). Birdsong and human speech: common themes and mechanisms. *Annual review of neuroscience*, 22(1):567–631.

Drake, C. and Botte, M.-C. (1993). Tempo sensitivity in auditory sequences: Evidence for a multiple-look model. *Perception & psychophysics*, 54(3):277–286.

Drayna, D., Manichaikul, A., Lange, M. d., Snieder, H., and Spector, T. (2001). Genetic correlates of musical pitch recognition in humans. *Science*, 291(5510):1969–1972.

Durrant, P. (2017). Lexical bundles and disciplinary variation in university students' writing: Mapping the territories. *Applied Linguistics*, 38(2):165–193.

Eccles, R., Van der Linde, J., le Roux, M., Holloway, J., MacCutcheon, D., Ljung, R., and Swanepoel, D. W. (2021). Is phonological awareness related to pitch, rhythm, and speech-in-noise discrimination in young children? *Language, Speech, and Hearing Services in Schools*, 52(1):383–395.

Eerola, T., Himberg, T., Toiviainen, P., and Louhivuori, J. (2006). Perceived complexity of western and african folk melodies by western and african listeners. *Psychology of Music*, 34(3):337–371.

Einhäuser, W. (2017). The pupil as marker of cognitive processes. In *Computational and cognitive neuroscience of vision*, pages 141–169. Springer.

Fabb, N. and Halle, M. (2012). Grouping in the stressing of words, in metrical verse, and in music. *Language and music as cognitive systems*, pages 4–21.

Facoetti, A., Lorusso, M. L., Paganoni, P., Cattaneo, C., Galli, R., Umilta, C., and Mascetti, G. G. (2003). Auditory and visual automatic attention deficits in developmental dyslexia. *Cognitive brain research*, 16(2):185–191.

Facoetti, A., Trussardi, A. N., Ruffino, M., Lorusso, M. L., Cattaneo, C., Galli, R., Molteni, M., and Zorzi, M. (2010). Multisensory spatial attention deficits are predictive of phonological decoding skills in developmental dyslexia. *Journal of cognitive neuroscience*, 22(5):1011–1025.

Falk, S. and Dalla Bella, S. (2016). It is better when expected: aligning speech and motor rhythms enhances verbal processing. *Language, Cognition and Neuroscience*, 31(5):699–708.

Falk, S., Müller, T., and Dalla Bella, S. (2015). Non-verbal sensorimotor timing deficits in children and adolescents who stutter. *Frontiers in Psychology*, 6:847.

Fernald, A., Swingley, D., and Pinto, J. P. (2001). When half a word is enough: Infants can recognize spoken words using partial phonetic information. *Child development*, 72(4):1003–1015.

Fery, M., Bernard, C., Thoret, E., Kronland-Martinet, R., and Ystad, S. (2021). Audio-tactile perception of roughness. In *CMMR2021, Music in the IA Era*, page 303.

Fitch, W. T. (2013). Rhythmic cognition in humans and animals: distinguishing meter and pulse perception. *Frontiers in systems neuroscience*, 7:68.

Fitch, W. T. (2014). Toward a computational framework for cognitive biology: Unifying approaches from cognitive neuroscience and comparative cognition. *Physics of life reviews*, 11(3):329–364.

Fitch, W. T. (2015). The biology and evolution of musical rhythm: an update. *Structures in the mind: Essays on language, music, and cognition in honor of Ray Jackendoff*, 293.

Fitzpatrick, P., Romero, V., Amaral, J. L., Duncan, A., Barnard, H., Richardson, M. J., and Schmidt, R. (2017). Evaluating the importance of social motor synchronization and motor skill for understanding autism. *Autism Research*, 10(10):1687–1699.

Fiveash, A., Bedoin, N., Gordon, R. L., and Tillmann, B. (2021). Processing rhythm in speech and music: Shared mechanisms and implications for developmental speech and language disorders. *Neuropsychology*, 35(8):771.

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

Flaten, E., Marshall, S. A., Dittrich, A., and Trainor, L. J. (2022). Evidence for top-down metre perception in infancy as shown by primed neural responses to an ambiguous rhythm. *European Journal of Neuroscience*, 55(8):2003–2023.

Flaugnacco, E., Lopez, L., Terribili, C., Montico, M., Zoia, S., and Schön, D. (2015). Music training increases phonological awareness and reading skills in developmental dyslexia: A randomized control trial. *PloS one*, 10(9):e0138715.

Forgeard, M., Schlaug, G., Norton, A., Rosam, C., Iyengar, U., and Winner, E. (2008). The relation between music and phonological processing in normal-reading children and children with dyslexia. *Music perception*, 25(4):383–390.

Fotidzis, T. S. (2020). *Phonology, Prosody, and Reading Skills: A Mismatch Negativity Experiment*. PhD thesis, Middle Tennessee State University.

Fotidzis, T. S., Moon, H., Steele, J. R., and Magne, C. L. (2018). Cross-modal priming effect of rhythm on visual word recognition and its relationships to music aptitude and reading achievement. *Brain sciences*, 8(12):210.

Fox, J. (2015). *Applied regression analysis and generalized linear models*. Sage Publications.

François, C., Chobert, J., Besson, M., and Schön, D. (2013). Music training for the development of speech segmentation. *Cerebral Cortex*, 23(9):2038–2043.

Franich, K., Wong, H. Y., Yu, A. C., and To, C. K. (2021). Temporal coordination and prosodic structure in autism spectrum disorder: timing across speech and non-speech motor domains. *Journal of Autism and Developmental Disorders*, 51(8):2929–2949.

Frey, A., François, C., Chobert, J., Besson, M., and Ziegler, J. C. (2019). Behavioral and electrophysiological investigation of speech perception deficits in silence, noise and envelope conditions in developmental dyslexia. *Neuropsychologia*, 130:3–12.

Friston, K. (2010). The free-energy principle: a unified brain theory? *Nature reviews neuroscience*, 11(2):127–138.

Fritzsche, T. and Höhle, B. (2015). Phonological and lexical mismatch detection in 30-month-olds and adults measured by pupillometry. In *ICPhS*.

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

Fujioka, T., Trainor, L. J., Large, E. W., and Ross, B. (2012). Internalized timing of isochronous sounds is represented in neuromagnetic beta oscillations. *Journal of Neuroscience*, 32(5):1791–1802.

Gallace, A. and Spence, C. (2014). *In touch with the future: The sense of touch from cognitive neuroscience to virtual reality*. OUP Oxford.

Geiser, E., Sandmann, P., Jäncke, L., and Meyer, M. (2010). Refinement of metre perception–training increases hierarchical metre processing. *European Journal of Neuroscience*, 32(11):1979–1985.

Gerken, L. and McGregor, K. (1998). An overview of prosody and its role in normal and disordered child language. *American Journal of Speech-Language Pathology*, 7(2):38–48.

Germagnoli, S. et al. (2016). La musica nella riabilitazione della dislessia: dati a favore dell'efficacia del training lettura ritmica. *La musica nella riabilitazione della dislessia: dati a favore dell'efficacia del training lettura ritmica*, pages 21–35.

Gervain, J. (2018). Gateway to language: The perception of prosody at birth. In *Boundaries crossed, at the interfaces of morphosyntax, phonology, pragmatics and semantics*, pages 373–384. Springer.

Gervain, J., Berent, I., and Werker, J. F. (2012). Binding at birth: The newborn brain detects identity relations and sequential position in speech. *Journal of Cognitive Neuroscience*, 24(3):564–574.

Gescheider, G. A. and Niblette, R. K. (1967). Cross-modality masking for touch and hearing. *Journal of experimental psychology*, 74(3):313.

Ghanizadeh, A. (2011). Sensory processing problems in children with adhd, a systematic review. *Psychiatry investigation*, 8(2):89.

Ghitza, O. (2011). Linking speech perception and neurophysiology: speech decoding guided by cascaded oscillators locked to the input rhythm. *Frontiers in psychology*, 2:130.

Gierhan, S. M. (2013). Connections for auditory language in the human brain. *Brain and language*, 127(2):205–221.

Giordano, M. (2016). *Vibrotactile feedback and stimulation in music performance*. McGill University (Canada).

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

Gordon, C. L., Cobb, P. R., and Balasubramaniam, R. (2018). Recruitment of the motor system during music listening: An ale meta-analysis of fmri data. *PloS one*, 13(11):e0207213.

Gordon, R. L., Fehd, H. M., and McCandliss, B. D. (2015). Does music training enhance literacy skills? a meta-analysis. *Frontiers in psychology*, 6:1777.

Gordon, R. L., Magne, C. L., and Large, E. W. (2011). Eeg correlates of song prosody: a new look at the relationship between linguistic and musical rhythm. *Frontiers in psychology*, 2:352.

Gordon, R. L., Ravignani, A., Hyland Bruno, J., Robinson, C. M., Scartozzi, A., Embalabala, R., Niarchou, M., 23andMe Research Team, Cox, N. J., and Creanza, N. (2021). Linking the genomic signatures of human beat synchronization and learned song in birds. *Philosophical Transactions of the Royal Society B*, 376(1835):20200329.

Goswami, U. (2002). Phonology, reading development, and dyslexia: A cross-linguistic perspective. *Annals of Dyslexia*, 52(1):139–163.

Goswami, U. (2011). A temporal sampling framework for developmental dyslexia. *Trends in cognitive sciences*, 15(1):3–10.

Goswami, U. and Bryant, P. (2016). *Phonological skills and learning to read*. Routledge.

Goswami, U., Gerson, D., and Astruc, L. (2010). Amplitude envelope perception, phonology and prosodic sensitivity in children with developmental dyslexia. *Reading and Writing*, 23(8):995–1019.

Goswami, U., Huss, M., Mead, N., Fosker, T., and Verney, J. P. (2013). Perception of patterns of musical beat distribution in phonological developmental dyslexia: Significant longitudinal relations with word reading and reading comprehension. *Cortex*, 49(5):1363–1376.

Gould, L., Mickleborough, M. J., Ekstrand, C., Lorentz, E., and Borowsky, R. (2017). Examining the neuroanatomical and the behavioural basis of the effect of basic rhythm on reading aloud. *Language, cognition and neuroscience*, 32(6):724–742.

Gould, L., Mickleborough, M. J., Lorentz, E., Ekstrand, C., and Borowsky, R. (2018). A behavioral and fmri examination of the effect of rhythm on reading noun-verb homographs aloud. *Language, cognition and neuroscience*, 33(7):829–849.

Grabe, E. and Low, E. L. (2002). Durational variability in speech and the rhythm class hypothesis. *Papers in laboratory phonology*, 7(1982):515–546.

Grafton, S. T., Hazeltine, E., and Ivry, R. (1995). Functional mapping of sequence learning in normal humans. *Journal of cognitive neuroscience*, 7(4):497–510.

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

Grahn, J. A. and Brett, M. (2007). Rhythm and beat perception in motor areas of the brain. *Journal of cognitive neuroscience*, 19(5):893–906.

Granier-Deferre, C. and Busnel, M.-C. (2011). L'audition prénatale, quoi de neuf? *Spirale*, (3):17–32.

Granier-Deferre, C., Ribeiro, A., Jacquet, A.-Y., and Bassereau, S. (2011). Near-term fetuses process temporal features of speech. *Developmental science*, 14(2):336–352.

Greenfield, M. D. and Roizen, I. (1993). Katydid synchronous chorusing is an evolutionarily stable outcome of female choice. *Nature*, 364(6438):618–620.

Greiffenstein, M. F., Baker, W. J., and Gola, T. (1994). Validation of malingered amnesia measures with a large clinical sample. *Psychological assessment*, 6(3):218.

Guenther, F. H. and Hickok, G. (2015). Role of the auditory system in speech production. *Handbook of clinical neurology*, 129:161–175.

György Buzsáki, M. (2019). *The brain from inside out*. Oxford University Press.

Hallé, P. A. and de Boysson-Bardies, B. (1996). The format of representation of recognized words in infants' early receptive lexicon. *Infant Behavior and Development*, 19(4):463–481.

Hämäläinen, J. A., Rupp, A., Soltész, F., Szücs, D., and Goswami, U. (2012). Reduced phase locking to slow amplitude modulation in adults with dyslexia: an meg study. *Neuroimage*, 59(3):2952–2961.

Hannon, E. E. and Johnson, S. P. (2005). Infants use meter to categorize rhythms and melodies: Implications for musical structure learning. *Cognitive psychology*, 50(4):354–377.

Hannon, E. E., Soley, G., and Ullal, S. (2012). Familiarity overrides complexity in rhythm perception: a cross-cultural comparison of american and turkish listeners. *Journal of Experimental Psychology: Human Perception and Performance*, 38(3):543.

Hannon, E. E. and Trainor, L. J. (2007). Music acquisition: effects of enculturation and formal training on development. *Trends in cognitive sciences*, 11(11):466–472.

Hannon, E. E. and Trehub, S. E. (2005). Tuning in to musical rhythms: Infants learn more readily than adults. *Proceedings of the National Academy of Sciences*, 102(35):12639–12643.

Hari, R. and Renvall, H. (2001). Impaired processing of rapid stimulus sequences in dyslexia. *Trends in cognitive sciences*, 5(12):525–532.

Harrington, D. L., Zimbelman, J. L., Hinton, S. C., and Rao, S. M. (2010). Neural modulation of temporal encoding, maintenance, and decision processes. *Cerebral Cortex*, 20(6):1274–1285.

Hartman, D. E. (2009). Wechsler adult intelligence scale iv (wais iv): return of the gold standard. *Applied neuropsychology*, 16(1):85–87.

Hasegawa, A., Okanoya, K., Hasegawa, T., and Seki, Y. (2011). Rhythmic synchronization tapping to an audio–visual metronome in budgerigars. *Scientific reports*, 1(1):1–8.

Haueisen, J. and Knösche, T. R. (2001). Involuntary motor activity in pianists evoked by music perception. *Journal of cognitive neuroscience*, 13(6):786–792.

Hawke, J. L., Olson, R. K., Willcut, E. G., Wadsworth, S. J., and DeFries, J. C. (2009). Gender ratios for reading difficulties. *Dyslexia*, 15(3):239–242.

Hawks, Z. W. and Constantino, J. N. (2020). Neuropsychiatric "comorbidity" as causal influence in autism. *Journal of the American Academy of Child & Adolescent Psychiatry*, 59(2):229–235.

Hayes, B. (1995). *Metrical stress theory: Principles and case studies*. University of Chicago Press.

Heinze, G., Wallisch, C., and Dunkler, D. (2018). Variable selection–a review and recommendations for the practicing statistician. *Biometrical journal*, 60(3):431–449.

Helmholtz, H. L. (1954). On the sensations of tone, translated by aj ellis.

Henry, M. J., Herrmann, B., and Obleser, J. (2014). Entrained neural oscillations in multiple frequency bands comodulate behavior. *Proceedings of the National Academy of Sciences*, 111(41):14935–14940.

Hepach, R. and Westermann, G. (2016). Pupillometry in infancy research. *Journal of Cognition and Development*, 17(3):359–377.

Herrera, L., Lorenzo, O., Defior, S., Fernandez-Smith, G., and Costa-Giomi, E. (2011). Effects of phonological and musical training on the reading readiness of native-and foreign-spanish-speaking children. *Psychology of Music*, 39(1):68–81.

Hickey, P. and Race, E. (2021). Riding the slow wave: Exploring the role of entrained low-frequency oscillations in memory formation. *Neuropsychologia*, 160:107962.

Hickok, G., Houde, J., and Rong, F. (2011). Sensorimotor integration in speech processing: computational basis and neural organization. *Neuron*, 69(3):407–422.

Hohle, B., Schmitz, M., Santelmann, L. M., and Weissenborn, J. (2006). The recognition of discontinuous verbal dependencies by german 19-month-olds: Evidence for lexical and structural influences on children's early processing capacities. *Language Learning and Development*, 2(4):277–300.

Hopyan, T., Dennis, M., Weksberg, R., and Cytrynbaum, C. (2001). Music skills and the expressive interpretation of music in children with williams-beuren syndrome: Pitch, rhythm, melodic imagery, phrasing, and musical affect. *Child Neuropsychology*, 7(1):42–53.

Horst, J. S. and Hout, M. C. (2016). The novel object and unusual name (noun) database: A collection of novel images for use in experimental research. *Behavior research methods*, 48(4):1393–1409.

Houston, D. M., Jusczyk, P. W., Kuijpers, C., Coolen, R., and Cutler, A. (2000). Cross-language word segmentation by 9-month-olds. *Psychonomic Bulletin & Review*, 7(3):504–509.

Humphrey, T. (1964). Some correlations between the appearance of human fetal reflexes and the development of the nervous system. *Progress in Brain Research*, 4:93–135.

Hunter, M. A. and Ames, E. W. (1988). A multifactor model of infant preferences for novel and familiar stimuli. *Advances in infancy research*.

Huron, D. (2006). Sweet anticipation.

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

Ito, Y., Shiramatsu, T. I., Ishida, N., Oshima, K., Magami, K., and Takahashi, H. (2022). Spontaneous beat synchronization in rats: Neural dynamics and motor entrainment. *Science Advances*, 8(45):eabo7019.

Iversen, J. R. and Patel, A. D. (2010). The beat alignment test (bat). In *International Conference for Music Perception and Cognition, Sapporo, Japan*.

Jackendoff, R. and Lerdahl, F. (2006). The capacity for music: What is it, and what's special about it? *Cognition*, 100(1):33–72.

Jackson, I. and Sirois, S. (2009). Infant cognition: going full factorial with pupil dilation. *Developmental science*, 12(4):670–679.

Jacobs, A. M. and Grainger, J. (1994). Models of visual word recognition: sampling the state of the art. *Journal of Experimental Psychology: Human perception and performance*, 20(6):1311.

Jacoby, W. G. (2000). Loess:: a nonparametric, graphical tool for depicting relationships between variables. *Electoral studies*, 19(4):577–613.

Jadoul, Y., Ravignani, A., Thompson, B., Filippi, P., and De Boer, B. (2016). Seeking temporal predictability in speech: comparing statistical approaches on 18 world languages. *Frontiers in human neuroscience*, 10:586.

Jentschke, S. and Koelsch, S. (2009). Musical training modulates the development of syntax processing in children. *Neuroimage*, 47(2):735–744.

Johnson, E. K. and Jusczyk, P. W. (2001). Word segmentation by 8-montholds: When speech cues count more than statistics. *Journal of memory and language*, 44(4):548–567.

Jones, M. R. (2018). *Time will tell: A theory of dynamic attending*. Oxford University Press.

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

Juntunen, M.-L. and Hyvönen, L. (2004). Embodiment in musical knowing: how body movement facilitates learning within dalcroze eurhythmics. *British Journal of Music Education*, 21(2):199–214.

Juntunen, T. and Hyvönen, A.-E. (2014). Resilience, security and the politics of processes. *Resilience*, 2(3):195–209.

Jusczyk, P. W. and Aslin, R. N. (1995). Infants detection of the sound patterns of words in fluent speech. *Cognitive psychology*, 29(1):1–23.

Jusczyk, P. W., Cutler, A., and Redanz, N. J. (1993). Infants' preference for the predominant stress patterns of english words. *Child development*, 64(3):675–687.

Jusczyk, P. W., Houston, D. M., and Newsome, M. (1999). The beginnings of word segmentation in english-learning infants. *Cognitive psychology*, 39(3-4):159–207.

Kadesjö, B. and Gillberg, C. (2001). The comorbidity of adhd in the general population of swedish school-age children. *Journal of child psychology and psychiatry*, 42(4):487–492.

Kalashnikova, M., Burnham, D., and Goswami, U. (2021). Rhythm discrimination and metronome tapping in 4-year-old children at risk for developmental dyslexia. *Cognitive Development*, 60:101129.

Kalashnikova, M., Goswami, U., and Burnham, D. (2019). Sensitivity to amplitude envelope rise time in infancy and vocabulary development at 3 years: A significant relationship. *Developmental science*, 22(6):e12836.

Kalender, B., Trehub, S. E., and Schellenberg, E. G. (2013). Cross-cultural differences in meter perception. *Psychological Research*, 77(2):196–203.

Kaminsky, S. K., Linsenmair, K. E., and Grafe, T. U. (1999). Reproductive timing, nest construction and tadpole guidance in the african pig-nosed frog, hemisus marmoratus. *Journal of Herpetology*, 33(1):119–123.

Karam, M., Nespoli, G., Russo, F., and Fels, D. I. (2009a). Modelling perceptual elements of music in a vibrotactile display for deaf users: A field study. In *2009 Second International Conferences on Advances in Computer-Human Interactions*, pages 249–254. IEEE.

Karam, M., Russo, F. A., and Fels, D. I. (2009b). Designing the model human cochlea: An ambient crossmodal audio-tactile display. *IEEE Transactions on Haptics*, 2(3):160–169.

Karatekin, C. (2007). Eye tracking studies of normative and atypical development. *Developmental review*, 27(3):283–348.

Karmiloff-Smith, A. (2007). Williams syndrome. *Current Biology*, 17(24):R1035–R1036.

Kaur, M., Srinivasan, S. M., and Bhat, A. N. (2018). Comparing motor performance, praxis, coordination, and interpersonal synchrony between children with and without autism spectrum disorder (asd). *Research in developmental disabilities*, 72:79–95.

Kertész, C. and Honbolygó, F. (2021). Tapping to music predicts literacy skills of first-grade children. *Frontiers in psychology*, page 4412.

Klein-Tasman, B. P., Phillips, K. D., Lord, C. E., Mervis, C. B., and Gallo, F. (2009). Overlap with the autism spectrum in young children with williams syndrome. *Journal of developmental and behavioral pediatrics: JDBP*, 30(4):289.

Klingner, J. (2010). *Measuring cognitive load during visual tasks by combining pupillometry and eye tracking*. Stanford University.

Koelsch, S., Fritz, T., v. Cramon, D. Y., Müller, K., and Friederici, A. D. (2006). Investigating emotion with music: an fmri study. *Human brain mapping*, 27(3):239–250.

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

Kösem, A. and Van Wassenhove, V. (2017). Distinct contributions of low-and high-frequency neural oscillations to speech comprehension. *Language, cognition and neuroscience*, 32(5):536–544.

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

Kraus, N. and Chandrasekaran, B. (2010). Music training for the development of auditory skills. *Nature reviews neuroscience*, 11(8):599–605.

Kroese, J. M., Hynd, G. W., Knight, D. F., Hiemenz, J. R., and Hall, J. (2000). Clinical appraisal of spelling ability and its relationship to phonemic awareness (blending, segmenting, elision, and reversal), phonological memory, and reading in reading disabled, adhd, and normal children. *Reading and Writing*, 13(1):105–131.

Kronbichler, M., Hutzler, F., and Wimmer, H. (2002). Dyslexia: Verbal impairments in the absence of magnocellular impairments. *Neuroreport*, 13(5):617–620.

Kuhl, P. K. (2004). Early language acquisition: cracking the speech code. *Nature reviews neuroscience*, 5(11):831–843.

Kuhl, P. K., Stevens, E., Hayashi, A., Deguchi, T., Kiritani, S., and Iverson, P. (2006). Infants show a facilitation effect for native language phonetic perception between 6 and 12 months. *Developmental science*, 9(2):F13–F21.

Kushki, A., Chau, T., and Anagnostou, E. (2011). Handwriting difficulties in children with autism spectrum disorders: A scoping review. *Journal of autism and developmental disorders*, 41(12):1706–1716.

Lacquaniti, F., Terzuolo, C., and Viviani, P. (1983). The law relating the kinematic and figural aspects of drawing movements. *Acta psychologica*, 54(1-3):115–130.

Ladányi, E., Lukács, Á., and Gervain, J. (2021). Does rhythmic priming improve grammatical processing in hungarian-speaking children with and without developmental language disorder? *Developmental science*, 24(6):e13112.

Ladányi, E., Persici, V., Fiveash, A., Tillmann, B., and Gordon, R. L. (2020). Is atypical rhythm a risk factor for developmental speech and language disorders? *Wiley Interdisciplinary Reviews: Cognitive Science*, 11(5):e1528.

Lahav, A., Saltzman, E., and Schlaug, G. (2007). Action representation of sound: audiomotor recognition network while listening to newly acquired actions. *Journal of Neuroscience*, 27(2):308–314.

Lam, H. M. Y. (2011). Assessment of preschoolers' gross motor proficiency: revisiting bruininks–oseretsky test of motor proficiency. *Early Child Development and Care*, 181(2):189–201.

Landa, R. J., Holman, K. C., O'Neill, A. H., and Stuart, E. A. (2011). Intervention targeting development of socially synchronous engagement in toddlers with autism spectrum disorder: A randomized controlled trial. *Journal of Child Psychology and Psychiatry*, 52(1):13–21.

Landerl, K. and Moll, K. (2010). Comorbidity of learning disorders: prevalence and familial transmission. *Journal of Child Psychology and Psychiatry*, 51(3):287–294.

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

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

Lanzilotti, C., Dumas, R., Grassi, M., and Schön, D. (2019). Prolonged exposure to highly rhythmic music affects brain dynamics and perception. *Neuropsychologia*, 129:191–199.

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

Large, E. W. and Snyder, J. S. (2009). Pulse and meter as neural resonance. *Annals of the New York Academy of Sciences*, 1169(1):46–57.

Larsson, M. (2012). Incidental sounds of locomotion in animal cognition. *Animal Cognition*, 15(1):1–13.

Larsson, M. (2015). Tool-use-associated sound in the evolution of language. *Animal cognition*, 18(5):993–1005.

Lashley, K. S. et al. (1951). *The problem of serial order in behavior*, volume 21. Bobbs-Merrill Oxford.

Lê, M., Quémart, P., Potocki, A., Gimenes, M., Chesnet, D., and Lambert, E. (2020). Rhythm in the blood: The influence of rhythm skills on literacy development in third graders. *Journal of Experimental Child Psychology*, 198:104880.

Lecanuet, J.-P. and Schaal, B. (2002). Sensory performances in the human foetus: A brief summary of research. *Intellectica*, 34(1):29–56.

Lee, H.-Y., Sie, Y.-S., Chen, S.-C., and Cheng, M.-C. (2015). The music perception performance of children with and without dyslexia in taiwan. *Psychological reports*, 116(1):13–22.

Lee, Y. S., Ahn, S., Holt, R. F., and Schellenberg, E. G. (2020). Rhythm and syntax processing in school-age children. *Developmental Psychology*, 56(9):1632.

Lehiste, I. (1970). Temporal organization of higher-level linguistic units. *The Journal of the Acoustical Society of America*, 48(1A):111–111.

Lense, M. D. and Dykens, E. M. (2016). Beat perception and sociability: Evidence from williams syndrome. *Frontiers in Psychology*, 7:886.

Lense, M. D., Ladányi, E., Rabinowitch, T.-C., Trainor, L., and Gordon, R. (2021). Rhythm and timing as vulnerabilities in neurodevelopmental disorders. *Philosophical Transactions of the Royal Society B*, 376(1835):20200327.

Leong, V. and Goswami, U. (2014). Impaired extraction of speech rhythm from temporal modulation patterns in speech in developmental dyslexia. *Frontiers in human neuroscience*, 8:96.

Lew-Williams, C., Ferguson, B., Abu-Zhaya, R., and Seidl, A. (2019). Social touch interacts with infants' learning of auditory patterns. *Developmental cognitive neuroscience*, 35:66–74.

Lewkowicz, D. J. (2014). Early experience and multisensory perceptual narrowing. *Developmental psychobiology*, 56(2):292–315.

Li, H., Helpard, L., Ekeroot, J., Rohani, S. A., Zhu, N., Rask-Andersen, H., Ladak, H. M., and Agrawal, S. (2021). Three-dimensional tonotopic mapping of the human cochlea based on synchrotron radiation phase-contrast imaging. *Scientific reports*, 11(1):1–8.

Lindeman, R. W., Yanagida, Y., Sibert, J. L., and Lavine, R. (2003). Effective vibrotactile cueing in a visual search task. In *Proc. of Interact 2003*, pages 89–96.

Lindstrom, M. J. and Bates, D. M. (1990). Nonlinear mixed effects models for repeated measures data. *Biometrics*, pages 673–687.

Liu, B., Wu, G., and Meng, X. (2012). Cross-modal priming effect based on short-term experience of ecologically unrelated audio-visual information: An event-related potential study. *Neuroscience*, 223:21–27.

Lizarazu, M., Lallier, M., Molinaro, N., Bourguignon, M., Paz-Alonso, P. M., Lerma-Usabiaga, G., and Carreiras, M. (2015). Developmental evaluation of atypical auditory sampling in dyslexia: Functional and structural evidence. *Human brain mapping*, 36(12):4986–5002.

Luo, C., Guo, Z.-w., Lai, Y.-x., Liao, W., Liu, Q., Kendrick, K. M., Yao, D.-z., and Li, H. (2012). Musical training induces functional plasticity in perceptual and motor networks: insights from resting-state fmri. *PLoS one*, 7(5):e36568.

Mackenzie, S. J., Getchell, N., Deutsch, K., Wilms-Floet, A., Clark, J. E., and Whitall, J. (2008). Multi-limb coordination and rhythmic variability under varying sensory availability conditions in children with dcd. *Human Movement Science*, 27(2):256–269.

Madison, G. and Merker, B. (2002). On the limits of anisochrony in pulse attribution. *Psychological research*, 66(3):201–207.

Mani, N., Mills, D. L., and Plunkett, K. (2012). Vowels in early words: An event-related potential study. *Developmental Science*, 15(1):2–11.

Mani, N. and Plunkett, K. (2011). Phonological priming and cohort effects in toddlers. *Cognition*, 121(2):196–206.

Manning, F. and Schutz, M. (2013). “moving to the beat” improves timing perception. *Psychonomic bulletin & review*, 20(6):1133–1139.

Marcus, G. F., Vijayan, S., Bandi Rao, S., and Vishton, P. M. (1999). Rule learning by seven-month-old infants. *Science*, 283(5398):77–80.

Mareva, S. and Holmes, J. (2019). Transdiagnostic associations across communication, cognitive, and behavioural problems in a developmentally at-risk population: A network approach. *BMC pediatrics*, 19(1):1–12.

Marks, L. E. (1979). Summation of vibrotactile intensity: An analog to auditory critical bands? *Sensory processes*.

Marsh, K. L., Isenhower, R. W., Richardson, M. J., Helt, M., Verbalis, A. D., Schmidt, R. C., and Fein, D. (2013). Autism and social disconnection in interpersonal rocking. *Frontiers in integrative neuroscience*, 7:4.

Martens, M. A., Reutens, D. C., and Wilson, S. J. (2010). Auditory cortical volumes and musical ability in williams syndrome. *Neuropsychologia*, 48(9):2602–2609.

Martens, M. A., Wilson, S. J., and Reutens, D. C. (2008). Research review: Williams syndrome: a critical review of the cognitive, behavioral, and neuroanatomical phenotype. *Journal of Child Psychology and Psychiatry*, 49(6):576–608.

Martin, R. M. (1975). Effects of familiar and complex stimuli on infant attention. *Developmental Psychology*, 11(2):178.

Martino, M. G., Pappalardo, F., Re, A. M., Tressoldi, P. E., Lucangeli, D., and Cornoldi, C. (2011). La valutazione della dislessia nell'adulto. *Un contributo alla standardizzazione della Batteria dell'Universita di Padova. Dislessia*, 8:119–134.

Mathôt, S. (2018). Pupillometry: Psychology, physiology, and function. *Journal of Cognition*, 1(1).

Mathôt, S., Schreij, D., and Theeuwes, J. (2012). Opensesame: An open-source, graphical experiment builder for the social sciences. *Behavior research methods*, 44(2):314–324.

Mathôt, S. and Van der Stigchel, S. (2015). New light on the mind's eye: The pupillary light response as active vision. *Current directions in psychological science*, 24(5):374–378.

Mathôt, S. and Vilotijević, A. (2022). Methods in cognitive pupillometry: Design, preprocessing, and statistical analysis. *bioRxiv*.

Maurer, D. and Werker, J. F. (2014). Perceptual narrowing during infancy: A comparison of language and faces. *Developmental psychobiology*, 56(2):154–178.

Mayall, L. A., D'Souza, H., Hill, E. L., Karmiloff-Smith, A., Tolmie, A., and Farran, E. K. (2021). Motor abilities and the motor profile in individuals with williams syndrome. *Advances in Neurodevelopmental Disorders*, 5(1):46–60.

Mayes, S. D., Calhoun, S. L., and Crowell, E. W. (2000). Learning disabilities and adhd: Overlapping spectrum disorders. *Journal of learning disabilities*, 33(5):417–424.

McAuley, J. D. (2010). Tempo and rhythm. In *Music perception*, pages 165–199. Springer.

McCloskey, M. and Rapp, B. (2017). Developmental dysgraphia: An overview and framework for research. *Cognitive neuropsychology*, 34(3-4):65–82.

Mehler, J., Jusczyk, P., Lambertz, G., Halsted, N., Bertoncini, J., and Amiel-Tison, C. (1988). A precursor of language acquisition in young infants. *Cognition*, 29(2):143–178.

Mekyska, J., Faundez-Zanuy, M., Mzourek, Z., Galaz, Z., Smekal, Z., and Rosenblum, S. (2016). Identification and rating of developmental dysgraphia by handwriting analysis. *IEEE Transactions on Human-Machine Systems*, 47(2):235–248.

Merchant, H. and Honing, H. (2014). Are non-human primates capable of rhythmic entrainment? evidence for the gradual audiomotor evolution hypothesis. *Frontiers in neuroscience*, page 274.

Merker, B. H., Madison, G. S., and Eckerdal, P. (2009). On the role and origin of isochrony in human rhythmic entrainment. *Cortex*, 45(1):4–17.

Miendlarzewska, E. A. and Trost, W. J. (2014). How musical training affects cognitive development: rhythm, reward and other modulating variables. *Frontiers in neuroscience*, page 279.

Mirabella, G., Del Signore, S., Lakens, D., Averna, R., Penge, R., and Capozzi, F. (2017). Developmental coordination disorder affects the processing of action-related verbs. *Frontiers in Human Neuroscience*, 10:661.

Molinaro, N., Lizarazu, M., Lallier, M., Bourguignon, M., and Carreiras, M. (2016). Out-of-synchrony speech entrainment in developmental dyslexia. *Human brain mapping*, 37(8):2767–2783.

Moreno, S., Marques, C., Santos, A., Santos, M., Castro, S. L., and Besson, M. (2009). Musical training influences linguistic abilities in 8-year-old children: more evidence for brain plasticity. *Cerebral cortex*, 19(3):712–723.

Morgan, J. L. and Saffran, J. R. (1995). Emerging integration of sequential and suprasegmental information in preverbal speech segmentation. *Child development*, 66(4):911–936.

Morillon, B. and Baillet, S. (2017). Motor origin of temporal predictions in auditory attention. *Proceedings of the National Academy of Sciences*, 114(42):E8913–E8921.

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

Mosing, M. A., Madison, G., Pedersen, N. L., Kuja-Halkola, R., and Ullén, F. (2014). Practice does not make perfect: no causal effect of music practice on music ability. *Psychological science*, 25(9):1795–1803.

Möttönen, R., Dutton, R., and Watkins, K. E. (2013). Auditory-motor processing of speech sounds. *Cerebral Cortex*, 23(5):1190–1197.

Motz, B. A., Erickson, M. A., and Hetrick, W. P. (2013). To the beat of your own drum: Cortical regularization of non-integer ratio rhythms toward metrical patterns. *Brain and cognition*, 81(3):329–336.

Munakata, Y. and Pfaffly, J. (2004). Hebbian learning and development. *Developmental science*, 7(2):141–148.

Muneaux, M., Ziegler, J. C., Truc, C., Thomson, J., and Goswami, U. (2004). Deficits in beat perception and dyslexia: Evidence from french. *NeuroReport*, 15(8):1255–1259.

Musacchia, G., Sams, M., Skoe, E., and Kraus, N. (2007). Musicians have enhanced subcortical auditory and audiovisual processing of speech and music. *Proceedings of the National Academy of Sciences*, 104(40):15894–15898.

Nayak, S., Coleman, P. L., Ladányi, E., Nitin, R., Gustavson, D. E., Fisher, S. E., Magne, C. L., and Gordon, R. L. (2022). The musical abilities, pleiotropy, language, and environment (maple) framework for understanding musicality-language links across the lifespan. *Neurobiology of Language*, 3(4):615–664.

Nazzi, T., Bertoncini, J., and Mehler, J. (1998). Language discrimination by newborns: toward an understanding of the role of rhythm. *Journal of Experimental Psychology: Human perception and performance*, 24(3):756.

Nazzi, T., Paterson, S., and Karmiloff-Smith, A. (2003). Early word segmentation by infants and toddlers with williams syndrome. *Infancy*, 4(2):251–271.

Nazzi, T. and Ramus, F. (2003). Perception and acquisition of linguistic rhythm by infants. *Speech communication*, 41(1):233–243.

Nespor, M., Shukla, M., and Mehler, J. (2011). Stress-timed vs. syllable-timed languages. *The Blackwell companion to phonology*, pages 1–13.

Nespor, M., Shukla, M., van de Vijver, R., Avesani, C., Schraudolf, H., and Donati, C. (2008). Different phrasal prominence realizations in vo and ov languages. *Lingue e linguaggio*, 7(2):139–168.

Nespor, M. and Vogel, I. (2007). Foreword to prosodic domains and external sandhi rules. *Annali Online di Ferrara-Lettere*, 1(14/53).

Neves, L., Correia, A. I., Castro, S. L., Martins, D., and Lima, C. F. (2022). Does music training enhance auditory and linguistic processing? a systematic review and meta-analysis of behavioral and brain evidence. *Neuroscience & Biobehavioral Reviews*, page 104777.

Niarchou, M., Gustavson, D. E., Sathirapongsasuti, J. F., Anglada-Tort, M., Eising, E., Bell, E., McArthur, E., Straub, P., McAuley, J. D., Capra, J. A., et al. (2021). Unravelling the genetic architecture of musical rhythm: a large-scale genome-wide association study of beat synchronization. *BioRxiv*, page 836197.

Nicolson, R. I. and Fawcett, A. J. (1990). Automaticity: A new framework for dyslexia research? *Cognition*, 35(2):159–182.

Nicolson, R. I. and Fawcett, A. J. (2007). Procedural learning difficulties: reuniting the developmental disorders? *TRENDS in Neurosciences*, 30(4):135–141.

Nicolson, R. I. and Fawcett, A. J. (2011). Dyslexia, dysgraphia, procedural learning and the cerebellum. *Cortex*, 47(1):117–127.

Nicolson, R. I., Fawcett, A. J., and Dean, P. (2001). Developmental dyslexia: the cerebellar deficit hypothesis. *Trends in neurosciences*, 24(9):508–511.

Nieuwenhuis, R., Te Grotenhuis, H., and Pelzer, B. (2012). Influence. me: tools for detecting influential data in mixed effects models.

Noreika, V., Falter, C. M., and Rubia, K. (2013). Timing deficits in attention-deficit/hyperactivity disorder (adhd): Evidence from neurocognitive and neuroimaging studies. *Neuropsychologia*, 51(2):235–266.

Nozaradan, S., Peretz, I., Missal, M., and Mouraux, A. (2011). Tagging the neuronal entrainment to beat and meter. *Journal of Neuroscience*, 31(28):10234–10240.

Nozaradan, S., Peretz, I., and Mouraux, A. (2012). Selective neuronal entrainment to the beat and meter embedded in a musical rhythm. *Journal of Neuroscience*, 32(49):17572–17581.

Oakes, L. M. (2017). Sample size, statistical power, and false conclusions in infant looking-time research. *Infancy*, 22(4):436–469.

Obermeier, C., Kotz, S. A., Jessen, S., Raettig, T., Von Koppenfels, M., and Menninghaus, W. (2016). Aesthetic appreciation of poetry correlates with ease of processing in event-related potentials. *Cognitive, Affective, & Behavioral Neuroscience*, 16(2):362–373.

Olander, L., Smith, A., and Zelaznik, H. N. (2010). Evidence that a motor timing deficit is a factor in the development of stuttering.

Ostiz-Blanco, M., Bernacer, J., Garcia-Arbizu, I., Diaz-Sanchez, P., Rello, L., Lallier, M., and Arrondo, G. (2021). Improving reading through videogames and digital apps: A systematic review. *Frontiers in psychology*, page 3835.

Overy, K. (2003). Dyslexia and music: From timing deficits to musical intervention. *Annals of the New York academy of sciences*, 999(1):497–505.

Ozernov-Palchik, O. and Patel, A. D. (2018). Musical rhythm and reading development: does beat processing matter? *Annals of the New York Academy of Sciences*, 1423(1):166–175.

Pagliarini, E., Guasti, M. T., Toneatto, C., Granocchio, E., Riva, F., Sarti, D., Molteni, B., and Stucchi, N. (2015). Dyslexic children fail to comply with the rhythmic constraints of handwriting. *Human Movement Science*, 42:161–182.

Pagliarini, E., Scocchia, L., Vernice, M., Zoppello, M., Balottin, U., Bouamama, S., Guasti, M. T., and Stucchi, N. (2017). Children's first handwriting productions show a rhythmic structure. *Scientific reports*, 7(1):1–10.

Pan, J., Klímová, M., McGuire, J. T., and Ling, S. (2022). Arousal-based pupil modulation is dictated by luminance. *Scientific reports*, 12(1):1–11.

Pasquini, E. S., Corriveau, K. H., and Goswami, U. (2007). Auditory processing of amplitude envelope rise time in adults diagnosed with developmental dyslexia. *Scientific Studies of Reading*, 11(3):259–286.

Patel, A. D. (2006). Musical rhythm, linguistic rhythm, and human evolution. *Music Perception*, 24(1):99–104.

Patel, A. D. (2011). Why would musical training benefit the neural encoding of speech? the opera hypothesis. *Frontiers in psychology*, 2:142.

Patel, A. D. (2021). Vocal learning as a preadaptation for the evolution of human beat perception and synchronization. *Philosophical Transactions of the Royal Society B*, 376(1835):20200326.

Patel, A. D. and Daniele, J. R. (2003). An empirical comparison of rhythm in language and music. *Cognition*, 87(1):B35–B45.

Patel, A. D. and Iversen, J. R. (2014). The evolutionary neuroscience of musical beat perception: the action simulation for auditory prediction (asap) hypothesis. *Frontiers in systems neuroscience*, 8:57.

Patel, A. D., Iversen, J. R., Bregman, M. R., and Schulz, I. (2009). Studying synchronization to a musical beat in nonhuman animals. *Annals of the New York Academy of Sciences*, 1169(1):459–469.

Patten, E., Watson, L. R., and Baranek, G. T. (2014). Temporal synchrony detection and associations with language in young children with asd. *Autism Research and Treatment*, 2014.

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

Penfield, W. and Rasmussen, T. (1950). The cerebral cortex of man; a clinical study of localization of function.

Peretz, I., Zatorre, R. J., et al. (2005). Brain organization for music processing. *Annual review of psychology*, 56(1):89–114.

Persici, V., Stucchi, N., and Arosio, F. (2019). Rhythmic and morphosyntactic predictions: The anticipation abilities of italian children with developmental dyslexia. In *BUCLD 43: Proceedings of the 43rd Annual Boston University Conference on Language Development*, pages 537–548.

Petitjean, C. (1989). *Une condition de l'audition foetale: la conduction sonore osseuse: conséquences cliniques et applications pratiques envisagées*. PhD thesis.

Petkov, C. I. and Jarvis, E. D. (2012). Birds, primates, and spoken language origins: behavioral phenotypes and neurobiological substrates. *Frontiers in evolutionary neuroscience*, 4:12.

Phillips-Silver, J. and Trainor, L. J. (2005). Feeling the beat: movement influences infant rhythm perception. *Science*, 308(5727):1430–1430.

Phillips-Silver, J. and Trainor, L. J. (2007). Hearing what the body feels: Auditory encoding of rhythmic movement. *Cognition*, 105(3):533–546.

Phillips-Silver, J. and Trainor, L. J. (2008). Vestibular influence on auditory metrical interpretation. *Brain and cognition*, 67(1):94–102.

Piek, J. P., Pitcher, T. M., and Hay, D. A. (1999). Motor coordination and kinaesthesis in boys with attention deficit–hyperactivity disorder. *Developmental Medicine and Child Neurology*, 41(3):159–165.

Pisoni, D. B., Kronenberger, W. G., Chandramouli, S. H., and Conway, C. M. (2016). Learning and memory processes following cochlear implantation: The missing piece of the puzzle. *Frontiers in psychology*, 7:493.

Poeppel, D. (2003). The analysis of speech in different temporal integration windows: cerebral lateralization as 'asymmetric sampling in time'. *Speech communication*, 41(1):245–255.

Poeppel, D. and Assaneo, M. F. (2020). Speech rhythms and their neural foundations. *Nature reviews neuroscience*, 21(6):322–334.

Politimou, N., Dalla Bella, S., Farrugia, N., and Franco, F. (2019). Born to speak and sing: Musical predictors of language development in preschoolers. *Frontiers in Psychology*, 10:948.

Polka, L., Sundara, M., and Blue, S. (2002). The role of language experience in word segmentation: A comparison of english, french, and bilingual infants. *The Journal of the Acoustical Society of America*, 111(5):2455–2455.

Provasi, J., Anderson, D. I., and Barbu-Roth, M. (2014). Rhythm perception, production, and synchronization during the perinatal period. *Frontiers in Psychology*, 5:1048.

Przybylski, L., Bedoin, N., Krifi-Papoz, S., Herbillon, V., Roch, D., Léculier, L., Kotz, S. A., and Tillmann, B. (2013). Rhythmic auditory stimulation influences syntactic processing in children with developmental language disorders. *Neuropsychology*, 27(1):121.

Puyjarinet, F., Bégel, V., Lopez, R., Dellacherie, D., and Dalla Bella, S. (2017). Children and adults with attention-deficit/hyperactivity disorder cannot move to the beat. *Scientific Reports*, 7(1):1–11.

R Core Team (2022). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Ramus, F., Nespor, M., and Mehler, J. (1999). Correlates of linguistic rhythm in the speech signal. *Cognition*, 73(3):265–292.

Ramus, F., Rosen, S., Dakin, S. C., Day, B. L., Castellote, J. M., White, S., and Frith, U. (2003). Theories of developmental dyslexia: insights from a multiple case study of dyslexic adults. *Brain*, 126(4):841–865.

Ravignani, A., Bowling, D. L., and Fitch, W. T. (2014). Chorusing, synchrony, and the evolutionary functions of rhythm. *Frontiers in psychology*, 5:1118.

Ravignani, A., Honing, H., and Kotz, S. A. (2017). The evolution of rhythm cognition: Timing in music and speech.

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

Ravignani, A. and Norton, P. (2017). Measuring rhythmic complexity: a primer to quantify and compare temporal structure in speech, movement, and animal vocalizations. *Journal of Language Evolution*, 2(1):4–19.

Reifinger Jr, J. L. (2019). Dyslexia in the music classroom: A review of literature. *Update: Applications of Research in Music Education*, 38(1):9–17.

Reisberg, D., Rappaport, I., and O'Shaughnessy, M. (1984). Limits of working memory: The digit digit-span. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 10(2):203.

Repp, B. H. (2005). Sensorimotor synchronization: A review of the tapping literature. *Psychonomic bulletin & review*, 12(6):969–992.

Repp, B. H. and Doggett, R. (2007). Tapping to a very slow beat: a comparison of musicians and nonmusicians. *Music Perception*, 24(4):367–376.

Repp, B. H. and Penel, A. (2004). Rhythmic movement is attracted more strongly to auditory than to visual rhythms. *Psychological research*, 68(4):252–270.

Richards, S. and Goswami, U. (2015). Auditory processing in specific language impairment (sli): Relations with the perception of lexical and phrasal stress. *Journal of Speech, Language, and Hearing Research*, 58(4):1292–1305.

Ritter, M., Colson, K. A., and Park, J. (2013). Reading intervention using interactive metronome in children with language and reading impairment: a preliminary investigation. *Communication Disorders Quarterly*, 34(2):106–119.

Rocha, S., Attaheri, A., Choisdealbha, Á. N., Brusini, P., Flanagan, S. A., Mead, N., Boutris, P., Gibbon, S., Olawole-Scott, H., Grey, C., et al. (2021). Infant sensorimotor synchronisation to speech and non-speech rhythms: A longitudinal study.

Roder, B. J., Bushnell, E. W., and Sasseville, A. M. (2000). Infants' preferences for familiarity and novelty during the course of visual processing. *Infancy*, 1(4):491–507.

Rosenblum, S., Parush, S., and Weiss, P. L. (2003). Computerized temporal handwriting characteristics of proficient and non-proficient handwriters. *The American Journal of Occupational Therapy*, 57(2):129–138.

Rueschemeyer, S.-A., Lindemann, O., van Rooij, D., van Dam, W., and Bekkering, H. (2010). Effects of intentional motor actions on embodied language processing. *Experimental Psychology*, 57(4):260.

Russo, S., Calignano, G., Dispaldro, M., and Valenza, E. (2021). An integrated perspective on spatio-temporal attention and infant language acquisition. *International journal of environmental research and public health*, 18(4):1592.

Russo, S. and Valenza, E. (2021). Apprendere attraverso il corpo, la musica e il ritmo: gli effetti positivi della sincronizzazione ritmica secondo i principi della cognizione incarnata. *Psicologia dell'educazione*, pages 23–40.

Sallat, S. and Jentschke, S. (2015). Music perception influences language acquisition: melodic and rhythmic-melodic perception in children with specific language impairment. *Behavioural Neurology*, 2015.

Sansavini, A., Bertoncini, J., and Giovanelli, G. (1997). Newborns discriminate the rhythm of multisyllabic stressed words. *Developmental Psychology*, 33(1):3.

Santolin, C., Russo, S., Calignano, G., Saffran, J. R., and Valenza, E. (2019). The role of prosody in infants' preference for speech: A comparison between speech and birdsong. *Infancy*, 24(5):827–833.

Sartori, G., Job, R., and Tressoldi, P. E. (1995). Batteria per la valutazione della dislessia e della disortografia evolutiva. *Firenze: Organizzazioni Speciali*.

Schachner, A., Brady, T. F., Pepperberg, I. M., and Hauser, M. D. (2009). Spontaneous motor entrainment to music in multiple vocal mimicking species. *Current Biology*, 19(10):831–836.

Schellenberg, E. G. (2011). Examining the association between music lessons and intelligence. *British journal of psychology*, 102(3):283–302.

Schlaug, G. (2011). Music, musicians, and brain plasticity. *The Oxford Handbook of Music Psychology*, pages 197–208.

Schmandt-Besserat, D. (2014). The evolution of writing. *International Encyclopedia of Social and Behavioral Sciences*, pages 1–15.

Schneider, T. R., Lorenz, S., Senkowski, D., and Engel, A. K. (2011). Gamma-band activity as a signature for cross-modal priming of auditory object recognition by active haptic exploration. *Journal of Neuroscience*, 31(7):2502–2510.

Schoemaker, M. M., Ketelaars, C. E., Van Zonneveld, M., Minderaa, R. B., and PhD, T. M. (2005). Deficits in motor control processes involved in production of graphic movements of children with attention-deficit-hyperactivity disorder. *Developmental Medicine & Child Neurology*, 47(6):390–395.

Schön, D. and Tillmann, B. (2015). Short-and long-term rhythmic interventions: perspectives for language rehabilitation. *Annals of the New York Academy of Sciences*, 1337(1):32–39.

Schroeder, C. E., Lakatos, P., Kajikawa, Y., Partan, S., and Puce, A. (2008). Neuronal oscillations and visual amplification of speech. *Trends in cognitive sciences*, 12(3):106–113.

Schwartze, M. and Kotz, S. A. (2013). A dual-pathway neural architecture for specific temporal prediction. *Neuroscience & Biobehavioral Reviews*, 37(10):2587–2596.

Seesjärvi, E., Särkämö, T., Vuoksimaa, E., Tervaniemi, M., Peretz, I., and Kaprio, J. (2016). The nature and nurture of melody: a twin study of musical pitch and rhythm perception. *Behavior genetics*, 46(4):506–515.

Seidl, A. and Cristià, A. (2008). Developmental changes in the weighting of prosodic cues. *Developmental Science*, 11(4):596–606.

Seidl, A., Tincoff, R., Baker, C., and Cristia, A. (2015). Why the body comes first: Effects of experimenter touch on infants' word finding. *Developmental science*, 18(1):155–164.

Seitz, J. A. (2005). Dalcroze, the body, movement and musicality. *Psychology of music*, 33(4):419–435.

Selkirk, E. (1984). On the major class features and syllable theory. *Language sound structure*.

Selkirk, E. (1995). Sentence prosody: Intonation, stress, and phrasing. *The handbook of phonological theory*, 1:550–569.

Shaywitz, S. E. and Shaywitz, B. A. (2005). Dyslexia (specific reading disability). *Biological psychiatry*, 57(11):1301–1309.

Silverman, A. K., Fey, N. P., Portillo, A., Walden, J. G., Bosker, G., and Neptune, R. R. (2008). Compensatory mechanisms in below-knee amputee gait in response to increasing steady-state walking speeds. *Gait & posture*, 28(4):602–609.

Sirois, S. and Brisson, J. (2014). Pupillometry. *Wiley Interdisciplinary Reviews: Cognitive Science*, 5(6):679–692.

Sirois, S. and Jackson, I. (2007). Pupil dilation and infant cognition. In *2007 IEEE 6th International Conference on Development and Learning*, pages 223–228. IEEE.

Snowling, M. J. (2013). Early identification and interventions for dyslexia: a contemporary view. *Journal of Research in Special Educational Needs*, 13(1):7–14.

Sohmer, H., Perez, R., Sichel, J.-Y., Priner, R., and Freeman, S. (2001). The pathway enabling external sounds to reach and excite the fetal inner ear. *Audiology and Neurotology*, 6(3):109–116.

Soley, G. and Hannon, E. E. (2010). Infants prefer the musical meter of their own culture: a cross-cultural comparison. *Developmental psychology*, 46(1):286.

Soltész, F., Szűcs, D., Leong, V., White, S., and Goswami, U. (2013). Differential entrainment of neuroelectric delta oscillations in developmental dyslexia. *PLoS One*, 8(10):e76608.

Srinivasan, S. M., Kaur, M., Park, I. K., Gifford, T. D., Marsh, K. L., and Bhat, A. N. (2015). The effects of rhythm and robotic interventions on the imitation/praxis, interpersonal synchrony, and motor performance of children with autism spectrum disorder (asd): a pilot randomized controlled trial. *Autism research and treatment*, 2015.

Stein, J. and Walsh, V. (1997). To see but not to read; the magnocellular theory of dyslexia. *Trends in neurosciences*, 20(4):147–152.

Stephan, M. A., Brown, R., Lega, C., and Penhune, V. (2016). Melodic priming of motor sequence performance: The role of the dorsal premotor cortex. *Frontiers in Neuroscience*, 10:210.

Strait, D. L., Parbery-Clark, A., O'Connell, S., and Kraus, N. (2013). Biological impact of preschool music classes on processing speech in noise. *Developmental cognitive neuroscience*, 6:51–60.

Stubley, E. (1999). Musical listening as bodily experience. *Canadian Journal of Research in Music Education*, 40(4):5–7.

Su, Y.-H. and Pöppel, E. (2012). Body movement enhances the extraction of temporal structures in auditory sequences. *Psychological research*, 76(3):373–382.

Sun, C., Meng, X., Du, B., Zhang, Y., Liu, L., Dong, Q., Georgiou, G. K., and Nan, Y. (2022). Behavioral and neural rhythm sensitivities predict phonological awareness and word reading development in chinese. *Brain and Language*, 230:105126.

Surányi, Z., Csépe, V., Richardson, U., Thomson, J. M., Honbolygó, F., and Goswami, U. (2009). Sensitivity to rhythmic parameters in dyslexic children: A comparison of hungarian and english. *Reading and Writing*, 22(1):41–56.

Swingley, D. and Aslin, R. N. (2000). Spoken word recognition and lexical representation in very young children. *Cognition*, 76(2):147–166.

Tallal, P. (1980). Auditory temporal perception, phonics, and reading disabilities in children. *Brain and language*, 9(2):182–198.

Tallal, P. (2004). Improving language and literacy is a matter of time. *Nature Reviews Neuroscience*, 5(9):721–728.

Tamási, K., McKean, C., Gafos, A., Fritzsche, T., and Höhle, B. (2017). Pupillometry registers toddlers' sensitivity to degrees of mispronunciation. *Journal of experimental child psychology*, 153:140–148.

Tamási, K., McKean, C., Gafos, A., and Höhle, B. (2019). Children's gradient sensitivity to phonological mismatch: considering the dynamics of looking behavior and pupil dilation. *Journal of Child Language*, 46(1):1–23.

Tanaka, Y., Kanakogi, Y., Kawasaki, M., and Myowa, M. (2018). The integration of audio- tactile information is modulated by multimodal social interaction with physical contact in infancy. *Developmental cognitive neuroscience*, 30:31–40.

Teie, D. (2016). A comparative analysis of the universal elements of music and the fetal environment. *Frontiers in Psychology*, 7:1158.

Ten Oever, S., Schroeder, C. E., Poeppel, D., Van Atteveldt, N., and Zion-Golumbic, E. (2014). Rhythmicity and cross-modal temporal cues facilitate detection. *Neuropsychologia*, 63:43–50.

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

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

Tichko, P., Kim, J. C., and Large, E. W. (2021). Bouncing the network: A dynamical systems model of auditory–vestibular interactions underlying infants' perception of musical rhythm. *Developmental Science*, 24(5):e13103.

Tichko, P., Kim, J. C., and Large, E. W. (2022). A dynamical, radically embodied, and ecological theory of rhythm development. *Songs and Signs: Interdisciplinary Perspectives on Cultural Transmission and Inheritance in Human and Nonhuman Animals*.

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

Tincoff, R., Seidl, A., Buckley, L., Wojcik, C., and Cristia, A. (2019). Feeling the way to words: Parents' speech and touch cues highlight word-to-world mappings of body parts. *Language learning and development*, 15(2):103–125.

Todd, N. P. and Lee, C. S. (2015). The sensory-motor theory of rhythm and beat induction 20 years on: a new synthesis and future perspectives. *Frontiers in human neuroscience*, 9:444.

Tomchek, S. D., Huebner, R. A., and Dunn, W. (2014). Patterns of sensory processing in children with an autism spectrum disorder. *Research in Autism Spectrum Disorders*, 8(9):1214–1224.

Toyomura, A., Fujii, T., and Kuriki, S. (2011). Effect of external auditory pacing on the neural activity of stuttering speakers. *Neuroimage*, 57(4):1507–1516.

Trainor, L. J., Chang, A., Cairney, J., and Li, Y.-C. (2018). Is auditory perceptual timing a core deficit of developmental coordination disorder? *Annals of the New York Academy of Sciences*, 1423(1):30–39.

Trainor, L. J., Gao, X., Lei, J.-j., Lehtovaara, K., and Harris, L. R. (2009). The primal role of the vestibular system in determining musical rhythm. *cortex*, 45(1):35–43.

Trainor, L. J. and Hannon, E. E. (2013). Musical development.

Trehub, S. E. (2003). The developmental origins of musicality. *Nature neuroscience*, 6(7):669–673.

Tressoldi, P., Cornoldi, C., and Re, A. (2013). Bvsco-2. *Batteria per la valutazione della scrittura e della competenza ortografica-2*.

Tuller, B. and Fowler, C. A. (1980). Some articulatory correlates of perceptual isochrony. *Perception & Psychophysics*, 27(4):277–283.

Tyack, P. L. (2020). A taxonomy for vocal learning. *Philosophical Transactions of the Royal Society B*, 375(1789):20180406.

Ullal-Gupta, S., Vanden Bosch der Nederlanden, C. M., Tichko, P., Lahav, A., and Hannon, E. E. (2013). Linking prenatal experience to the emerging musical mind. *Frontiers in systems neuroscience*, 7:48.

Valera, E. M., Spencer, R. M., Zeffiro, T. A., Makris, N., Spencer, T. J., Faraone, S. V., Biederman, J., and Seidman, L. J. (2010). Neural substrates of impaired sensorimotor timing in adult attention-deficit/hyperactivity disorder. *Biological psychiatry*, 68(4):359–367.

Van Gemmert, A. W. and Van Galen, G. P. (1997). Stress, neuromotor noise, and human performance: a theoretical perspective. *Journal of Experimental Psychology: Human Perception and Performance*, 23(5):1299.

van Rij, J., Hendriks, P., van Rijn, H., Baayen, R. H., and Wood, S. N. (2019). Analyzing the time course of pupillometric data. *Trends in hearing*, 23:2331216519832483.

Velleman, S. L. and Vihman, M. M. (2007). Phonology in infancy and early childhood: implications for theories of language learning. In *Phonology in context*, pages 25–50. Springer.

Vihman, M. M., Nakai, S., DePaolis, R. A., and Hallé, P. (2004). The role of accentual pattern in early lexical representation. *Journal of Memory and Language*, 50(3):336–353.

Vinegard, M. (1994). A revised adult dyslexia check list. *EDUCARE-LONDON-NATIONAL BUREAU FOR HANDICAPPED STUDENTS-*, pages 21–21.

Viviani, P. and Terzuolo, C. (1982). Trajectory determines movement dynamics. *Neuroscience*, 7(2):431–437.

Von Bekesy, G. (1959). Similarities between hearing and skin sensations. *Psychological review*, 66(1):1.

von Békésy, G. and Peake, W. T. (1990). Experiments in hearing.

Vouloumanos, A. and Werker, J. F. (2004). Tuned to the signal: the privileged status of speech for young infants. *Developmental science*, 7(3):270–276.

Vuolo, J., Goffman, L., and Zelaznik, H. N. (2017). Deficits in coordinative bimanual timing precision in children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, 60(2):393–405.

Wagenmakers, E.-J. and Farrell, S. (2004). Aic model selection using akaike weights. *Psychonomic bulletin & review*, 11(1):192–196.

Wallin, N. L., Merker, B., and Brown, S. (2001). *The origins of music*. MIT press.

Wan, C. Y., Rüber, T., Hohmann, A., and Schlaug, G. (2010). The therapeutic effects of singing in neurological disorders. *Music perception*, 27(4):287–295.

White, K. S. and Morgan, J. L. (2008). Sub-segmental detail in early lexical representations. *Journal of Memory and Language*, 59(1):114–132.

White, K. S., Yee, E., Blumstein, S. E., and Morgan, J. L. (2013). Adults show less sensitivity to phonetic detail in unfamiliar words, too. *Journal of Memory and Language*, 68(4):362–378.

White, L., Payne, E., and Mattys, S. L. (2009). Rhythmic and prosodic contrast in venetan and sicilian italian. *Phonetics and phonology: Interactions and interrelations*, pages 137–158.

Wickham, H. (2016). Programming with ggplot2. In *ggplot2*, pages 241–253. Springer.

Wieland, E. A., McAuley, J. D., Dilley, L. C., and Chang, S.-E. (2015). Evidence for a rhythm perception deficit in children who stutter. *Brain and language*, 144:26–34.

Wilson, M. and Cook, P. F. (2016). Rhythmic entrainment: why humans want to, fireflies can't help it, pet birds try, and sea lions have to be bribed. *Psychonomic bulletin & review*, 23(6):1647–1659.

Wilson, P. H., Smits-Engelsman, B., Caeyenberghs, K., Steenbergen, B., Sugden, D., Clark, J., Mumford, N., and Blank, R. (2017). Cognitive and neuroimaging findings in developmental coordination disorder: new insights from a systematic review of recent research. *Developmental Medicine & Child Neurology*, 59(11):1117–1129.

Winkler, I., Háden, G. P., Ladinig, O., Sziller, I., and Honing, H. (2009). Newborn infants detect the beat in music. *Proceedings of the National Academy of Sciences*, 106(7):2468–2471.

Wolff, P. H. (2002). Timing precision and rhythm in developmental dyslexia. *Reading and Writing*, 15(1):179–206.

Wood, S. and Part (2007). R mgcv-package.

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

Yairi, E. and Ambrose, N. (2013). Epidemiology of stuttering: 21st century advances. *Journal of fluency disorders*, 38(2):66–87.

Yates, C. M., Justus, T., Atalay, N. B., Mert, N., and Trehub, S. E. (2017). Effects of musical training and culture on meter perception. *Psychology of Music*, 45(2):231–245.

Yoshida, K. A., Fennell, C. T., Swingley, D., and Werker, J. F. (2009). Fourteen-month-old infants learn similar-sounding words. *Developmental science*, 12(3):412–418.

Yu, C., Suanda, S. H., and Smith, L. B. (2019). Infant sustained attention but not joint attention to objects at 9 months predicts vocabulary at 12 and 15 months. *Developmental science*, 22(1):e12735.

Zatorre, R. J., Chen, J. L., and Penhune, V. B. (2007). When the brain plays music: auditory–motor interactions in music perception and production. *Nature reviews neuroscience*, 8(7):547–558.

Zelaznik, H. N. and Goffman, L. (2010). Generalized motor abilities and timing behavior in children with specific language impairment.

Zhang, J. D., Susino, M., McPherson, G. E., and Schubert, E. (2020). The definition of a musician in music psychology: A literature review and the six-year rule. *Psychology of Music*, 48(3):389–409.

Zhang, N. and Zhang, Q. (2019). Rhythmic pattern facilitates speech production: An erp study. *Scientific reports*, 9(1):1–11.

Zhao, T. C. and Kuhl, P. K. (2016). Musical intervention enhances infants' neural processing of temporal structure in music and speech. *Proceedings of the National Academy of Sciences*, 113(19):5212–5217.

Zimmerman, E. and Barlow, S. M. (2012). The effects of vestibular stimulation rate and magnitude of acceleration on central pattern generation for chest wall kinematics in preterm infants. *Journal of Perinatology*, 32(8):614–620.