UNIVERSITY OF PADOVA

Department of General Psychology

Master Degree in Cognitive Neuorscience and Clinical Neuropsychology

Final dissertation

The Impact of Prenatal Language Environment on Newborns: A Meta-Analysis of fNIRS Studies

Supervisor
Professor Judit Gervain

Co-supervisor
PhD Jessica Gemignani

Candidate: **Jessica Audrey Sweatman**
Student ID number: **2094640**

Academic Year 2023-2024

Acknowledgements

I would like to express my deepest gratitude to everyone who has been part of this journey and provided me with invaluable support. With the encouragement and guidance of those around me, I was able to finish this degree on time, despite starting one semester later than everyone else.

First and foremost, I would like to extend my sincere thanks to Professor Judit Gervain for your expert insights and constructive feedback. Your passion for developmental research in language acquisition and expertise in fNIRS has inspired me to write my thesis on this topic. I am incredibly grateful for the time and effort you dedicate to guiding me through this process.

I am immensely grateful to Doctor Jessica Gemignani for your support and mentorship. Your guidance and knowledge on meta-analysis helped me tremendously. Without your code and expertise in coding, I would have never been able to do a meta-analysis.

I would also like to extend my heartfelt appreciation to my parents, Simone and James Sweatman, who have been my anchor throughout this journey. Your unconditional love, understanding, and belief in me provided the strength I needed to persevere.

Lastly, to my friends and partner, thank you for being by my side and helping me maintain my sanity during challenging times. Your emotional support has been invaluable, and I cherish all the moments we shared together at university. Thank you all for believing in me and making this accomplishment possible.

Abstract

This meta-analysis synthesizes findings from functional Near-Infrared Spectroscopy (fNIRS) studies on neonates to investigate the neural mechanisms underlying language acquisition. The review focuses on studies that explore how prenatal language exposure might influence the development of the auditory system in fetuses by comparing the native language to a non-native language or a non-language stimulus. A total of eight peer-reviewed studies were selected for the literature review, and three studies were included in the meta-analysis based on predefined inclusion criteria. During the meta-analysis, MATLAB and JASP were used to combine and analyze the results. The findings revealed that the left temporal cortex was active during forward speech, suggesting early lateralization of language processing. Differences in neural activation were observed when neonates were exposed to languages of different rhythmic classes and backward speech. The results of this meta-analysis contribute to our understanding of how prenatal experiences influence language acquisition and auditory development.

Key words: fNIRS, prenatal language acquisition, neonates, language lateralization patterns, auditory cortex.

Table of Contents

1 Introduction

Understanding the neural mechanisms underlying language acquisition is a critical area in language acquisition research, as it sheds light on the debate of which abilities are innate or learned through experience. Researchers question whether brain specialization for language is an innate ability or whether language exposure shapes neural development. Advances in neuroimaging techniques such as functional near-infrared spectroscopy (fNIRS), have provided new opportunities to explore how infants perceive language in utero, by testing them right after birth before they are exposed to language outside of the womb. This meta-analysis aims to synthesize fNIRS studies investigating newborns' responses to the prenatally heard language vs. unfamiliar languages, providing a clearer understanding of early language discrimination and the effects of prenatal auditory exposure on brain development.

Since Broca's discovery of the importance of the left hemisphere in language perception and production, extensive research has aimed to pinpoint the exact brain areas and networks responsible for the different aspects of language processing. Most studies focused on adult participants, which established that the temporal lobe of the left hemisphere is specialized for language processing in most left-handed adults (Vigneau et al., 2006). However, the question remains whether this specialization is genetically determined or a result of exposure to one's native language. Noam Chomsky theorized that an innate ability is required to acquire language from the input of the external world. Empiricists theorize the input of the external world greatly influences language acquisition. Studying infants offers insights into the acquisition of these abilities during critical periods of language development (Nallet & Gervain, 2021).

The auditory system starts functioning around the twenty-fourth to the twenty-eighth week of gestation, exposing the fetus to auditory stimuli, which includes language spoken by the mother (Nallet & Gervain, 2021). The auditory stimuli heard in the womb passes through the mother's skin, tissue, and embryotic fluid before reaching the fetus (Abrams & Gerhardt, 1996). Additionally, the fetus can access the mother's speech via bone conduction (Abrams & Gerhardt, 1996). These factors attenuate the speech signal, suppressing a lot of its high-frequency content (Gerhardt & Abrams, 2000). As a result, the fetus can mostly only hear the low-frequency components of speech within the womb (Hepper & Shahidullah, 1994). This means that individual speech sounds are suppressed or greatly reduced, while speech prosody, such as rhythm and melody is preserved (Vogelsang et al., 2023).

Research suggests that infants can hear and remember what they have heard in the womb. For instance, they prefer speech over sine waves, their mother's voice over others, and human speech over monkey calls (Vouloumanos & Werker,2004; Kisilevsky et al., 2009; Vouloumanos et al., 2010). One of the hypotheses is that infants learn prosodic patterns in utero as a bootstrapping mechanism which will later aid in language acquisition (Gervain, 2018). The developing auditory system in neonates allows them to already distinguish between languages from different rhythmic classes as observed in behavioral studies, by measuring sucking intensity and looking time, and neuroimaging studies, such as fNIRS and fMRI (Gasparini et al., 2021).

Earlier research primarily used behavioral methods such as sucking intensity, looking time, and heart rate, but as technology advances neuroimaging techniques such as fMRI and fNIRS are now being used even with the youngest infants (Nallet & Gervain, 2021). Compared to alternative neuroimaging techniques, fNIRS is better when testing young participants such as children and infants (Gervain et al., 2008). fNIRS is a neuroimaging technique that uses red and near-infrared light to investigate the neural activity of the cerebral cortex by measuring the changes in oxygenated and deoxygenated hemoglobin (Nallet & Gervain 2021). In other words, it does not measure neural activity directly, but rather indirectly, through its hemodynamic correlation. Its portability and ease of application make it ideal for infant research, offering precise spatial localization without the need for paste or gel (Arimitsu et al., 2018). fNIRS is applied to the participant's head via a cap with electrodes attached and can be used concurrently with EEG, as it does not interfere with data collection (Nallet & Gervain, 2021). Although fNIRS has lower temporal resolution compared to EEG, the two can be used together to provide comprehensive spatial and temporal data (Nallet & Gervain, 2021). Given these advantages, this thesis focuses on fNIRS studies, excluding all other neuroimaging techniques.

Human language is a complex auditory stimulus researchers have been trying hard to understand, including how and when we acquire the ability to process language. A few days after birth, neonates can differentiate between linguistic and non-linguistic sounds (Arimitsu et al., 2018). Infants prefer listening to speech to other non-speech stimuli, such as sine waves and backward speech (Vouloumanos & Werker, 2004). Understanding prosodic patterns and the rhythmic class difference is a fundamental step for language acquisition as rhythmicity is associated with word order in language (Gervain et al., 2008).

Neonates' ability to discriminate languages of different rhythmic classes, regardless of familiarity, is well-supported by behavioral and neuroimaging studies (Gasparini et al., 2021). fNIRS can further elucidate the brain regions involved in this discrimination (Nallet & Gervain, 2021). Human languages can be classified within the three rhythmic classes by timing and stress patterns of speech (Ramus et al., 1999). The language class can be divided based on two parameters, which are the percent vowel duration within a sequence and the standard deviation of the duration of consonantal intervals (Gasparini et al., 2021). The three classes include mora-timed, stress-timed, and syllable-timed.

The studies included in this thesis do not use the same languages within their stimuli, but most of them compare a different rhythmic class language to the participant's native language. Japanese is a mora-timed language (Sato et al., 2012). English and Arabic are stress-timed languages (Vannesing et al., 2016; Sato et al., 2012). Syllable-timed languages include Tagalog, Italian, Finnish, Spanish, and French (May et al., 2011; May et al., 2017; Pena et al., 2003; Kotilahti et al., 2009). Since the ability to discriminate between different rhythmic classed languages has already been established, the authors of most newborn NIRS language discrimination studies investigate if the brain specialization for language is a universal ability, such as the ability to discriminate different rhythmic class languages, or if it is influenced by prenatal language experience (May et al., 2011).

The first of the newborn NIRS studies investigating the neural mechanisms of speech perception and language discrimination conducted by Pena et al. (2003) did not use an alternative rhythmic class as another stimulus. Inspired by behavioral studies suggesting infants were able to distinguish between forward but not backward speech, Pena et al. (2003), used fNIRS to investigate the neural mechanisms. Pena et al. (2003) focused solely on the neonate's native language, Italian, and compared it to a non-language stimulus, backward speech. The unique aspect of backward speech is that it cannot be reproduced by the human vocal tract, but has the same duration, pitch, and intensity as forward speech allowing it to carry the same acoustic properties (Pena et al., 2003). The inclusion of backward speech has been adopted in all subsequent experiments. Pena et al.'s (2003) study was revolutionary as it found infants' brain areas on the left hemisphere were active during forward speech, but not during backward speech, similar to how adults' brains would be active during normal speech.

After Pena et al. (2003)'s findings, many studies were conducted to understand the extent to which language is processed in the neonatal brain by using fNIRS. The key question multiple studies asked is how the prosodic information obtained through a low pass frequency filter impacts the development of language acquisition by looking at the activity of the brain. The studies explored how the prenatal experience of the infant's mother language can aid in language acquisition by comparing different rhythmic class languages. Although the results of the studies differ slightly, native language is always found to activate the left anterior temporal areas of the neonates' brain, which is similar to what is found in adult brains. The conflicting conclusions of these studies lie in the results found with the non-native language. Some studies found that the left hemisphere was active regardless of familiarity with the languages, while other studies found that the right hemisphere was active when the language was unfamiliar. Bilateral activation towards an unfamiliar language is found in adults (Olulade et al., 2020). Are infants' brains left lateralized to all languages and lose this ability as they acquire their first language, or do infants need to acquire their first language for speech to be left-lateralized?

An additional two studies investigated the fine line between language and non-language stimuli in neonates (May et al., 2017; Kotilahti et al., 2009). Left hemisphere specialization is only seen in speech and not in non-linguistic signals such as scrambled speech, sine-wave contours, tones, monkey calls, and backward speech in young infants (Dehaene-Lambertz et al., 2002; Minagawa-Kawai et al., 2011; Peña et al., 2003; Perani et al., 2011; Shultz et al., 2014; Taga et al., 2007). As a neonate's auditory system is still developing and has not been exposed to the full band of speech, would the left hemisphere be activated by prosodically similar stimuli such as music and whistled speech?

To what extent will an auditory sound be perceived as language? In the second study of May et al. (2017), the authors investigated the extent to which neonates perceive language by using a Silbo Gomero, a whistled language originating from Spanish, as the non-language stimulus. Will Silbo be processed as a spoken language and activate the same brain areas as familiar and non-familiar languages, or will it be processed as a non-linguistic sound stimulus? Silbo is a unique form of communication used in the Canary Islands to communicate from long distances and unlike normal language, it is no one's first acquired language (Carreiras et al., 2005). In adults who know and practice Silbo, the surrogate language activates the same brain regions as familiar normal speech (Carreiras et al., 2005). Silbo is a surrogate language and does not fall under one of the rhythmic classes, but as it derives from Spanish it maintains similar structure, rhythm, prosody, and communicative intent. While Silbo is produced in an alternative way compared to spoken language, it has a limited phonetic repertoire and a reduced acoustic complexity. May et al. (2017) used fNIRS to measure the brain activity in native English neonates while they listened to Spanish and Silbo, forward and backward.

Kotilahti et al. (2009) wanted to understand the neural activity difference between classical music and spoken language. Music and language share similar properties such as rhythm, the presence of rule-based structures, prosody, and pitch (Kotilahti et al., 2009). Previous research found that infants prefer songs that have been played to them while they are in the womb (Gerhardt & Abrams, 2000). By using fNIRS, researchers have tried to understand the impact of prenatal language exposure, but also to what extent neonates perceive music and language as two separate sound stimuli (Kotilahti et al., 2009).

Understanding the neural mechanisms underlying language acquisition is crucial in exploring whether language abilities are innate or shaped by experience. Before birth, the fetus is exposed to low-frequency sounds allowing the fetus to learn the prosodic patterns of their native language. Knowing the prosodic patterns of their native language may aid the neonate in language acquisition in the early stages of life (Gervain, 2018). Due to previous studies confirming that neonates can discriminate between two different rhythmic-based languages, fNIRS is done to understand how the brain processes these different languages. Do they have the ability to process all languages equally, or does the exposure in utero give them a head start in learning their native language? This thesis aims to compile a literature review on a collection of fNIRS studies on neonates' ability to distinguish their native language from other rhythmically classed languages based on the prosodic information they received while in utero, as well as include a brief meta-analysis on 3 of the studies.

2. Theoretical Background

How we acquire our first language has long fascinated psychologists and linguists. Various theories, ranging from innatism to empiricism, attempt to explain this process, but the truth may lie somewhere in between. Our first language is acquired through a combination of exposure to our native language and innate mechanisms active during the critical periods of development. Newborns, who have not yet experienced language outside the womb, provide an ideal opportunity to explore nature's role in language acquisition. However, previous research suggests that the fetus is exposed to sounds at a low frequency, which may be the first step in language acquisition. Due to modern day advancements in technology, it is possible to study neonates in an ethical and noninvasive way.

2.1 Nature vs Nurture Debate

There is an ongoing debate regarding the extent to which genetics or the environment play a role in our development. Conducting research on young children and infants allows researchers to determine which abilities are innate and which ones are gained through exposure as the young children and infants have little experience. It is theorized that lack of an ability during the early stages of life would suggest this ability is formed by the environment rather than an innate ability.

2.1.1 Behavioralist Theory

The behavioralist theory of language acquisition, also known as the imitation theory, was proposed by B.F. Skinner to explain language acquisition, focusing heavily on the impact of the environment

(Skinner, 1969). Behavioralist theory states that language is learned mainly through imitation and external motivation is the driving force for language acquisition (Skinner, 1957). The child learns their first language by imitating their caregivers or parents and then modifying their language through operant conditioning (Skinner, 1969). When a child mimics their caregiver and uses the correct phrase or word, they are positively reinforced through praise or given the reward they were asking for, such as food, a toy, etc. Negative reinforcement occurs when the child uses the wrong phrase or word and is corrected, scolded, or ignored by the caregiver. Through positive and negative reinforcements, the child learns the correct way to use the words and phrases they have been exposed to by their caregivers (Skinner, 1969).

2.1.2 Constructivist Theory

The constructivist theory, proposed by Jean Piaget, suggests that the environment and mental abilities both contribute to language acquisition in infants (Piaget, 1952). Language acquisition occurs within the context of the child's cognitive or mental development (Piaget, 1955). Children build schemas of the external world as their language develops with new experiences (Piaget, 1955). As a child develops, they understand concepts and then express them through language (1955). Both Piaget and Skinner put emphasis on the importance of environmental stimuli driving language acquisition, but Piaget suggests there is a greater importance of an innate mechanism.

2.1.3 Nativist Theory

Noam Chomsky's Nativist Theory suggests that humans are biologically predisposed to acquire language, meaning that language acquisition does not depend exclusively on environmental conditioning (Chomsky, 1959). Infants are born with an innate mechanism, known as the Language Acquisition Device (LAD), which enables them to acquire language efficiently and rapidly regardless of the richness or quality of the environment (Chomsky, 1965). The LAD equips children with the ability to figure out and understand the structure of the language they're exposed to and generate grammar rules based on limited input.

Chomsky's "poverty of stimulus" argument suggests that the linguistic input children receive from their environment is too limited on its own to account for their ability to grasp the full complexity of their native language (Chomsky, 1986). Essentially, exposure alone is insufficient for language acquisition and therefore an internal mental process would need to occur. Additionally, language acquisition through exposure is impossible due to the induction problem. The induction problem states that for any finite data set, such as the language samples children hear, there is an infinite number of potential rule systems they could theoretically infer. Despite this, children develop a correct understanding of their native language's grammar, which would suggest the presence of an innate linguistic mechanism.

Moreover, humans use language productively, meaning they can create an infinite variety of unique sentences and phrases beyond what they have heard. This ability indicates that language acquisition is not merely a process of mimicry, but rather one of internalizing and applying grammatical rules (Chomsky, 1959). In conclusion, the nativist theory claims that the human brain is hardwired for language learning, which universal principles governing how we acquire language (Chomsky, 1986).

2.1.4 Nature AND Nurture

Empiricists such as behavioralists and to some extent constructivists argue that the quality and type of linguistic input holds more value than Chomsky initially proposed (Skinner, 1957). Non-linguistic cues, such as physical objects and interacting with a social partner can play a key role in aiding young children to acquire a language (Chomsky, 1986). The quality of the linguistic cues, like the types of words the caregivers use and how they address infants and children through infant directed speech (IDS), are also incredibly crucial IDS, characterized by exaggerated pitch and slower tempo, helps infants understand and process language by drawing their attention to the meaning behind the speech.

In addition to these external factors, cognitive abilities and general learning mechanisms may contribute to how infants acquire language. These include skills like making analogies, developing Theory of Mind (the ability to understand others' thoughts and feelings), and joint attention, where both the child and caregiver focus on the same object or event. Together, these cognitive and social factors highlight the importance of interaction and mental processes in the development of language. Rather than viewing it as nature versus nurture, researchers have shifted their interests to see how nature and nurture work together to bring about development (Choi et al., 2018; Gervain & Mehler, 2010; Werker & Tees, 2005).

2.2 Development of the Auditory System

The debate between nature and nurture remains unresolved, as the human auditory system begins developing in utero and becomes functional between the 24th and 28th week of gestation (Eggermont & Moore, 2012). This prevents research from investigating the true initial state of language learning, i.e. a state when an infant does not yet have language experience at all. After this gestational age, the fetus is capable of hearing and is exposed to language even before birth. Animal studies suggest that what is heard in the uterus is not the same as what is heard outside the uterus (Abrams & Gerhardt, 1996). Sounds, including spoken language, get distorted and low-pass filtered (approximately at 400-600Hz) through the maternal tissue and amniotic fluid (Gerhardt & Abrams, 2000). This low-pass filtering maintains theprosodic information of speech. It is theorized that the fetus learns the prosodic patterns of their maternal language, which may help them outside the womb with language acquisition (Gervain, 2018). This prenatal period may begin shaping the specialization of language processing in the left hemisphere, much like it does in adults.

The adult human brain exhibits bilateral activation when processing a foreign language, while activity tends to be more lateralized to the left hemisphere when listening to one's native language (Jeong et al., 2007). Since infants are born linguistic citizens of the world, would there be lateralization towards any language (or sounds that are similar to language) and overtime as the ability to natively acquire a language diminishes, our brains become specialized to only our native language? Or is the neonatal brain developing to be left lateralized toward their native language based on the low frequency speech they heard within utero and over time the asymmetry becomes stronger (Silver et al., 2021)? The current thesis investigates this question.

2.3 Prosodic Bootstrapping Theory

Several theories and hypotheses try to explain how we acquire our native language, with the prosodic bootstrapping theory highlighting the importance of prenatal experience in laying the foundation for later language acquisition (Gervain, 2018). As previously mentioned, the fetus' environment in the womb allows them to hear speech through a low pass filter (Abrams & Gerhardt, 1996). This filter allows the fetus to exclusively hear prosodic information, while limiting the full spectrum of speech to enter the womb (Gerhardt & Abrams 2000). This gradual exposure to prosodic features of language is believed to provide an early advantage, by introducing the fetus to fundamental elements of speech before they encounter the full complexity of language (Gervain, 2018).

Although exposure to prosodic cues in the womb is not necessary for later language acquisition as can be seen in studies involving premature infants, infants to deaf parents, and deaf infants, it offers infants an advantage to later language acquisition (Rago et al., 2014; Arimitsu et al. 2018). Learning the prosodic patterns of their native language prenatally enables them to more easily process speech postnatally (Nallet & Gervain, 2021). By recognizing these patterns, neonates are able to segment speech into more manageable units (Nallet & Gervain, 2018). Immediately after birth, neonates can differentiate between languages of different rhythmic class (Ramus et al., 1999). This is because infants use prosodic cues to recognize rhythmic patterns for segmenting speech and building linguistic structure (Gasparini et al., 2021). Prosodic bootstrapping serves as the first step in breaking down speech and eventually understanding the syntactic structure of their native language (Gervain, 2018).

2.4 Rhythmic Class Discrimination

With a functioning but not fully mature auditory system, newborns are sensitive to linguistic contrasts and process the ability to acquire any human language (Eggermont & Moore, 2012). Neonates can discriminate between their native language and a foreign language only if these languages fall under a different rhythmic class (Ramus et al., 1999) This is because they learned the prosodic patterns of their native language while in utero (Ramus et al., 1999). Rhythmic classes are defined by how syllables and stresses shape the rhythm of speech, influencing both how it is produced and how it is perceived (Ramus et al., 1999). These patterns affect the timing and flow of spoken language, aiding listeners in distinguishing between various types of rhythmic structures (Nazzi et al., 2019).

The initial studies focused on a limited set of eight languages, but over time, various metrices have been developed to categorize all languages into the three distinct rhythmic classes (Gasparini et al., 2021). These three categories are mora-timed, stress-timed, and syllable-timed (Ramus et al., 1999). Although rhythmic class theory is criticized as an empirical way to group languages in phonetic literature, it has been corroborated through multiple studies to be reliable in developmental research, especially in explaining which languages infants can discriminate between (Hohle, et al., 2020).

2.4.1 Rhythmic Class Acquisition Hypothesis

Early in development, infants tend to rely more on rhythmic information than on segmental information, because prosodic cues alone are sufficient to help them to discriminate between rhythmically different languages (Molnar, Gervain, & Carreiras, 2013). This preference can be explained by the Rhythmic Class Acquisition Hypothesis, which posits that infants initially rely on the rhythmic patterns to segment speech and group languages (Nazzi et al., 2003). Neonates are able to distinguish between two languages belonging to different rhythmic classes but struggle to differentiate between languages originating from the same rhythmic class (Dorn et al., 2019). This hypothesis stems from the prosodic bootstrapping theory, which suggests that prosodic cues allow neonates to identify the rhythmic patterns of spoken language (Nazzi et al., 1998). But how are languages categorized into different rhythmic classes?

2.4.2 Native Language Acquisition Hypothesis

The Native Language Acquisition Hypothesis states that infants are biologically predisposed to acquire their native language, but over time, they become specialized in the language they're exposed to, gradually the ability to discriminate phonetic contrasts from non-native languages (Nazzi et al., 2000). Newborns are sensitive to rhythmic differences, as rhythm helps them detect boundaries between words and syllables. However, as they are increasingly exposed to only their native language, they fine-tune their abilities to align with that language's specific characteristics (Dorn et al., 2019). This means that while infants initially possess a broad sensitivity to multiple languages, they lose this ability, in favor of specializing in their native language (Dorn et al., 2019).

As infants gain familiarity with their native language, their reliance on rhythm sensitivity diminishes, allowing them to focus on more complex aspects of their native language, such as phonetic and grammatical structures (Butler et al., 2011). There is evidence of this phenomenon can be seen occurring in bilingual children, who fine-tune their abilities to both of their native languages (Mercure et al., 2020). Over time, these specialized abilities allow infants to master the unique linguistic features of their environment.

2.4.3 Metrices Involved in Differentiating All Languages into Rhythmic Classes

The initial study done by Ramus et al. (1999) used 8 languages to theorize the importance of rhythmic class. Since then, multiple studies have expanded from the initial 8 languages. A meta-analysis conducted by Gasparini et al. (2021) gathered studies that identified rhythmic class in 14 different languages and multiple accents. Regardless of their native language, infants could discriminate languages of different rhythmic classes, but not languages of the same class. Regarding this, how does one determine which language falls into which rhythmic class?

Since the initial study, the metrices used to determine the rhythmic class of a language have expanded to ensure the validity of the rhythmic class claim. Ramus et al. (1999) suggested that rhythmic class can be explained through the vocalic proportion (%V) and the variability of duration in consonants and vowels (ΔC and ΔV). Grabe and Low (2002) introduced the Pairwise Variability Indices as a new way to measure the rhythm of speech, which focuses on how the length of the sound changes from one to the next. They claim that this measure is more reliable as it considers differences between speakers and utterances (Grabe and Low, 2002). White and Mattys (2007) proposed a metric designed to take the speed of speech into account, VarcroV, as it measures the variability in vowel timing. As it takes into consideration the speed of the speaker, it provides a more accurate assessment of rhythmic patterns (White and Mattys, 2007). There are many more metrices that can be used to determine the rhythmic class of a language, but these are the most common ones used in developmental research (Gasparini et al., 2021).

2.4.4 Rhythmic Classification of Languages

With these metrices in mind, here are how the languages included in the thesis are categorized by rhythmic class. The three rhythmic categories are mora-timed, stressed-timed, and syllable-timed (Ramus et al., 1999). Although the categorization may seem straight forward for some languages, other languages lay at the cusp of two categories (Grabe & Low, 2002).

Mora-timed languages use mora, the unit of sound that determines syllable weight (Grabe & Low, 2002). Each mora is perceived with roughly equal duration. Japanese is a mora-timed language as the rhythm is organized around mora rather than syllables or stress patterns (Grabe & Lowe, 2002). In Japanese, a sentence like "Kon-ni-chi-wa", illustrates how each mora has a similar duration. In the study conducted by Sato et al. (2012) compare Japanese, a mora-timed language, to English, a stress-timed language.

The rhythm perceived in stress-timed languages is from the intervals between stressed syllables (Kazuo Ueda et al., 2017). This means that the rhythm is perceived as a series of beats occurring at relatively equal intervals. The unstressed syllables could be stretched or compressed to fit the timing of the stressed syllables (Grabe & Lowe, 2002). English and Arabic are stress-timed languages (Ramus et al., 1999). In English, a sentence like "The CAT chased the MOUSE" the capital letters are the stressed syllables, and the lowercase letters are the unstressed syllables. The timing between the stressed syllables is the same, while the unstress syllables are compressed, causes an uneven rhythm (Grabe & Lowe, 2002).

Syllable-timed languages use syllables as a rhythmic unit (Ramus et al., 2000). This means the rhythm is more even, and that each syllable is perceived to take approximately the same amount of time (Nazzi & Cutler, 2019). There is less variation in the length of syllables compared to stress-timed languages (Nazzi & Cutler, 2019). Languages that fall under the syllable-timed category include Italian, Spanish, Tagalog, Finnish, and French (Gasparini et al., 2021). There is some dispute regarding whether Tagalog and Finnish are syllable-timed languages, but a majority of studies state that these languages are syllable-timed (Gasparini et al., 2021). In Italian, a sentence like “Mi pia-ce la pi-zza”, illustrates how each syllable is roughly the same length in duration.

The studies used different syllable-timed languages, while English is used for a stress-timed rhythmic language. Pena et al. (2003) did not compare languages as they only used Italian, a syllable-timed language. In Kotilahti et al. (2009)’s study, Finnish was used as both the native language and the language stimulus. May et al. (2011) used native English neonates and compared their native language, a stress-timed language, to Tagalog, a syllable-timed language. In Vannassing et al. (2016) native French neonates listened to their native language, a syllable-timed language, and Arabic, a stress-timed language. The first study in May et al. (2017) used native English neonates to compare Spanish, a syllable-timed language, and English, a stress-timed language. However, the language used in Bartha-Doering et al. (2019) remains unknown.

2.5 Non-Language Stimuli

The first task of a newborn is to discriminate their native language from other sounds in their environment. However, this can be challenging as the category of speech is broad for neonates and young infants. Ongoing research aims to understand the extent to which various auditory stimuli, including monkey calls, sine waves, and music, may be perceived as speech. While backwards directed speech has never been considered a possible candidate for a language stimulus, it serves as a valuable comparison for assessing the boundaries of potential language perception (Saberi &

Perrott, 1999). An interesting area of study is surrogate languages, as adults fluent in a surrogate language display left hemisphere brain lateralization, similar to that seen in natural language (Carreiras et al., 2005). Unfortunately, surrogate languages cannot be the first language acquired, as they lack the complexity required for early language development (Carreiras et al., 2005). In this section the thesis explores the extent to which neonates perceive non-language stimuli as language.

2.5.1 Animal Calls

Humans and non-human primates can make identical sounds because our vocal cords share similar biological structures (Ekström, 2024). Both have vocal folds in the larynx that vibrate when air is passed through them to produce sound (Ekström, 2024). Due to the similarities between primates and humans, there are some studies that investigate whether neonates and older infants perceive monkey calls similar to how they would perceive human speech.

Newborn to 2-months-old show equal preference to human and Rhesus monkey calls, but at 3 months and older they prefer only the human voice (Vouloumanos et al., 2010; Minagawa-Kawai et al., 2011). Suggesting that infants initially categorize both human and non-human primate vocalizations as speech-like but fine-tune their preferences over time as they begin to specialize in processing human speech (Vouloumanos et al., 2010). This is particularly interesting when considering that monkey calls fall within the broader range of primate vocalizations, which infants might initially perceive as speech (Vouloumanos et al., 2010). Non-human primates lack the articulation and vocal control needed for speech, which results in short utterances rather than words and sentences (Ekström, 2024). In this study, Vouloumanos et al. (2010) used short utterances rather than full sentences, raising the possibility that longer, more complex vocalizations might elicit different results.

2.5.2 Backwards Directed Speech

Humans can mimic monkey calls as they fall under their vocal capabilities, but backwards directed speech cannot be produced by humans (Pena et al., 2003). Backwards speech is created by editing a sentence of normal speech spoken by a human in reverse. Although initially created by a human vocal cord, once reversed certain sounds are impossible for humans to replicate (Saberi & Perrot, 1999). This is because the phonetic structures are disrupted in reverse speech and the articulation required to produce the backwards phonemes are not in the human vocal cord range (Saberi & Perrott, 1999). Studies confirm that regardless of age, humans do not perceive backwards directed speech as language (Mushtaq et al., 2019). In neuroimaging studies, the brain regions activated during backward directed speech are not the same as the ones activated in forward directed speech (normal speech) (Pena et al., 2003).

Backward directed speech shares similar characteristics to forward directed speech, while not being perceived as a linguistic stimulus, making it an excellent control compared to another stimulus (Pena et al., 2003). The intensity and pitch of forward directed speech are still intact in backward directed speech (Saberi & Perrott, 1999). In studies, native speakers would read a script or story, and this recording would be used as the forward directed speech and after reversing the audio would also be the backwards directed speech (Pena et al.,2003). This means that the same recording of the native speaker’s voice would be used, but on one condition it will be modified preventing it from being perceived as a linguistic stimulus.

The pioneering study by Pena et al. (2003) used backwards directed speech as a control to investigate whether neonates' auditory cortex would activate in the same areas as an adult brain when listening to their native language (Italian). As compared to the backwards directed speech, forward directed speech showed activation in the left anterior temporal area (Pena et al., 2003). This was a remarkable finding as it suggested the developing auditory cortex of a neonate, who has yet to be exposed to language outside the womb, already processes language in the left hemisphere (Pena et al., 2003). Later researchers expanded this initial study to compare the neonate's native language to that of a different rhythmic class, but they continued to use backward directed speech as a control. There were mixed results on whether all language is processed in the left hemisphere in neonates, but all results pertaining to the cortical activity related to backwards speech are the same. Backwards directed speech, regardless of initial language will always activate both hemispheres of the brain like all non-linguistic sounds (Sato et al., 2011; May et al., 2017; Vannassing et al., 2016).

2.5.3 Music

A sound that has rhythm but is not defined in a rhythmic class is music. Although music is a non-linguistic sound (when it is played by an instrument), music is limited to certain rules in a similar manner that language is (Temperley, 2022). Music and language share many characteristics such as structure, processing and cognitive function (Rimas & Rimas, 2024). Music and language share some properties crucial for conveying meaning, such as pitch, melody, rhythm, and syntax (Temperley, 2022). Pitch is the frequency of a sound, and in language it is used in prosody to convey meaning of words (Temperley, 2022). In music, melody is the combination of pitches in a sequence. Rhythm is timing and pattern of sounds. In language, rhythm is presented in syllable stress patterns, word order, and sentence flow (Temperley, 2022). Classical music has complex rhythm structures to convey tension, resolution and emotional impact. Syntax in language refers to the rules and structure that govern how words are combined to form meaningful sentences. In music, particularly classical music, there is a form of musical syntax, where notes and chords are organized according to rules of harmony and composition (Termperley, 2022). These overlapping features may suggest that the neural circuits involved in processing music may be the same as the ones occurring in language processing.

While there are multiple adult studies that suggest music and language learning are closely related, there is disagreement on whether language and music are processed the same in the brain. Some studies suggest that music and language activate the same areas in adult musicians (Vuust et al., 2022). Other studies disprove this by testing aphasic patients maintain their musical abilities, suggesting that music and language neural networks are closely related, but not the same (Chen et al., 2023). In the auditory cortex of a neonate, these neural networks might not be fully developed, and it begs the question of whether neonates perceive language and music equally.

Kotilahti et al. (2010) investigated neonates brain activity response to classical music and the Finnish language. As mentioned previously, the Finnish language falls under the syllable-timed rhythmic class. Samples of the fairytale, the Little Red Ridding Hood were read by a Finnish voice actress in infant directed speech (Kotilahti et al., 2010). Compared to adult directed speech, infant directed speech has high variations in the intensity, pitch, and temporal structures (Kotilahti et al., 2010). The music was from a piano concert by Mozart. The results show no lateralization when newborns listened to music and only slight lateralization when listening to speech (Kotilahti et al., 2010). The slight lateralization was a group average, some neonates showed no lateralized differences between Finnish and classical music (Kotilahti, et al., 2010). These findings suggest that classical music does not fall under the broad category of speech for newborns. However, music is a vast and diverse category, encompassing various genres and styles, each with unique methods of composition and production. Perhaps humming, whistling, or drums would elicit a linguistic response in the neonatal brain.

2.5.4 Whistle Language

Surrogate languages replicate the phonological aspects of spoken language using non-verbal forms to communicate over long distances (Carreiras et al., 2005). These languages aren't always spoken but whistled, hummed, or created through drums (James, 2021). Unlike the non-human primate calls, surrogate languages do not use the larynx or vocal cords when communicating (James, 2021). Surrogate languages can be found around the world in different continents and used to communicate in different ways (James, 2021). There is not one type of a certain surrogate language. There are different types of whistle languages, such as the bird language in Turkey or the Silbo Gomera used in the Canary Islands (Gaynudtdinova & Mutallimova, 2021; González et al., 2020). They are both produced by whistling, but they are not the same.

Adults fluent in their surrogate language have lateral brain activation, while adults unfamiliar with this surrogate language will show bilateral brain activation (Carreiras et al., 2005). This means that surrogate languages are processed like real languages (González et al., 2020). Despite it being like a real language, a surrogate language cannot be acquired as a first language (Carreiras et al., 2005). This might be due to the fact that a surrogate language is built on the foundation of a "real" language (James, 2021). The structure of the surrogate language depends on the cultural aspects of the people and what the initial reason the surrogate language was invented for. For example, the whistle language Silbo Gomero was created to communicate between islands and originates from the Spanish language (Carreiras et al., 2005). Certain characteristics are the same between Silbo Gomero and Spanish (Carreiras et al., 2005). In this section we will focus exclusively on Silbo Gomero.

Since Silbo Gomero activates the language processing areas in the cerebral cortex in adults, would neonates' language category be broad enough to process Silbo Gomero as a language? May et al. (2017) investigate this in an fNIRS study by measuring native English neonates brain activity to compare Silbo Gomero and Spanish, the language it is based off of. At first glance one would assume it would make be better to use Spanish neonates to compare Silbo Gomero and Spanish (their native language), but there are many benefits as to why English infants were used. English neonates were a good choice as they were not prenatally exposed to Spanish or Silbo Gomero, and therefore both languages were novel. Spanish and English fall under different rhythmic classes (as mentioned earlier), but a surrogate language cannot be easily classified into a rhythmic class. This means neonates will be able to differentiate between the three languages.

The first study in May et al. (2017) was a baseline to see forward English and Spanish compared to backward English and Spanish. The results showed that English forward had a stronger lateral activation in the left hemisphere as compared to Spanish (May et al., 2017). In the second study, investigating Spanish and Silbo Gomero, forward and backward, Spanish showed stronger lateralized activation in the left hemisphere than Silbo Gomero which was bilateral activation (May et al., 2017). Compared to the first and second study, forward Spanish had a stronger lateralized when compared to Silbo Gomero than English (May et al., 2017). This may be because neonates tend to favor their native language (the one they were exposed to prenatally) over non-native languages, directing their attention to familiar sounds. However, in the absence of their native language, infants will shift their focus and begin processing other languages (May et al., 2017).

2.6 fNIRS

As technology has advanced over the years, so has neuroimaging techniques. Developmental research, initially relying on behavioral studies, can now investigate the developing brain through accurate but safe and non-invasive means (Gervain et al., 2011). Out of all the neuroimaging techniques, functional near-infrared spectroscopy (fNIRS) has many characteristics that make it an ideal tool for studying infants and young children (Gervain et al., 2011). This thesis will only focus on fNIRS data in the analysis.

2.6.1 The Fundamentals of fNIRS

Functional near-infrared spectroscopy (fNIRS) is a non-invasive neuroimaging technique that uses near-infrared light to measure hemodynamic changes in the brain (Pinti et al., 2018). Portable and easy to apply, the fNIRS cap is suitable for use across diverse populations as it can be worn by adults, children, and young infants (Wilcox & Biondi, 2015). The fNIRS device comprises a cap fitted with light sources and detectors (Gervain et al., 2011). Once securely on the head, the cap emits near-infrared light (650–950 nm) penetrating the scalp and underlying brain tissue (Pinti et al., 2018). This light interacts with the hemoglobin in the brain, which is either absorbed or reflected (Gervain et al., 2011). Absorption depends on the amount of oxygenated or deoxygenated hemoglobin in the blood (Pinti et al.,2018).

The cap's sensors capture the reflected light that returns to the scalp (Pinti et al., 2018). By measuring how much light is reflected or absorbed, the fNIRS system can calculate the relative concentrations of oxygenated and deoxygenated hemoglobin over time (Pinti et al., 2018). Since neural activity is accompanied by an increase in oxygen, this leads to an influx of oxygenated blood (Pinti et al., 2018). The more activity in the brain, the more oxygenated hemoglobin will be measured. This means fNIRS is an indirect measurement of brain activity (Pinti et al., 2018).

fNIRS offers excellent temporal resolution, capturing rapid changes in blood oxygenation over 1 to 2 seconds, making it ideal for monitoring brain function in real-time (Pinti et al., 2018). The increase in oxygenated hemoglobin detected by the system corresponds to heightened neural activity, allowing researchers to infer which brain regions are actively engaged during various cognitive or sensory tasks (Pinti et al., 2018). However, its spatial resolution is limited compared to functional magnetic resonance imaging (fMRI), as it primarily measures the cortex and not deep brain structures (Pinti et al., 2018). fNIRS measures the cortical surface with a spatial resolution of two to three centimeters (Pinti et al., 2018). Although it is impossible to obtain structural images and anatomical information (Pinti et al., 2018). Despite this limitation, fNIRS remains a valuable neuroimaging tool for research and clinical applications.

2.6.2 Advantages of fNIRS in Developmental Research

As a safe and durable neuroimaging technique, fNIRS is particularly well-suited for developmental research (Gervain et al., 2011). Unlike other neuroimaging techniques that use harmful radiation, deep brain stimulation, or a magnetic field, fNIRS uses light to measure brain activity on the surface of the brain (Wilcox & Biondi, 2015). Although fMRI and fNIRS use oxygenated hemoglobin to indirectly measure neural activation, fMRI can be uncomfortable, causing participants with claustrophobia anxiety (Aslin et al., 2015). The fNIRS cap is more comfortable compared to other neuroimaging tools and operates silently, which minimizes distractions (Pinti et al., 2018). The fNIRS device can accommodate significant movement and does not require participants to remain still, which is wonderful for restless children and infants too young to follow instructions (Wilcox & Biondi, 2015). The device is portable and flexible, allowing it to be set up in any environment, and can cater to infants sitting in their parents' lap or in a bassinet (Pinti et al., 2018).

As compared to other neuroimaging methods, the advantages of fNIRS are suitable for developmental research and outweigh the disadvantages. The limitation of only measuring the cortical surface of the brain doesn't necessarily have to be a limitation if the area of interest is only the cerebral cortex. The area of interest of this thesis is in the auditory cortex, and more specifically the language areas within the auditory cortex. Not all studies need to investigate the whole brain, as this would be both time-consuming and costly. Compared to fMRI and PET (positron emission tomography), fNIRS is more affordable and easier to implement, making it accessible to research teams with a lower budget (Penti et al., 2018). The fNIRS cap comes in multiple sizes and can adjust to any head size, allowing longitudinal or cross-sectional studies to use it in different age ranges (Wilcox & Biondi, 2015). fNIRS can be integrated with behavioral studies, allowing researchers to link behaviors and brain areas together (Wilcox & Biondi, 2015). EEG and fNIRS use the same cap during an experiment, as these two neuroimaging techniques do not interfere with one another (Pinti et al., 2018). Although the data collected from EEG and fNIRS can be integrated, it is extremely challenging as I will mention in the next section.

2.6.3 Reason to Only Include fNIRS

Researching children and infants can be challenging due to their limited attention span, inability to understand instructions, lack of interest, physical or motor limitations, and other factors (Wilcox & Biondi, 2015). Neonates are particularly difficult to study. Before neuroimaging techniques became more popular within the developmental research community, sucking rates and looking times were primarily used to assess young infants, as they were unable to communicate with the experimenter in any other way (Pinti et al., 2018). Although these experimental paradigms are reliable and have led to significant discoveries and theories, this thesis will focus on neuroimaging techniques, specifically fNIRS.

As mentioned earlier, EEG and fNIRS data can be merged, but it is challenging due to the differences in the procedures to obtain the data. fNIRS is an indirect measure of brain activity, measuring the oxygenated hemoglobin in the blood, while EEG is a direct measure, measuring the electric current produced by neurons in the brain (Li et al., 2022). EGG has a higher temporal resolution, but a lower spatial resolution as compared to fNIRS (Pinti et al., 2018). fNIRS has a lower temporal resolution due to the delay in the indirect measurement of brain activity compared to EEG (Li et al., 2022). EEG has a lower spatial resolution compared to fNIRS, which can target specific cortical regions (Pinti et al., 2018). With a appropriate methodology, combining the two neuroimaging techniques can be done (Li et al., 2022), but due to these challenges, EEG studies were not included within this meta-analysis.

3 Methods

3.1 Literature Search Strategy

A literature search was conducted to identify studies relevant to the meta-analysis on prenatal language acquisition in neonates using fNIRS. Google Scholar was used as a database as it is known for their extensive coverage of peer-reviewed studies in neuroscience and developmental psychology. The following keywords and combinations employed were: “fNIRS”, “neonates”, “native language” “unfamiliar language”, and “prenatal language acquisition”. The search was performed on 25/02/2024. The initial search on Google Scholar yielded a total of 5.030 articles.

3.2 Study Selection Process

Of the 5030 articles, non-experimental studies, i.e. reviews, meta-analyses etc., were excluded, yielding 1478 articles. Further selection was done by checking the following criteria:

1. Newborn participants between 1-5 days.

   Finding suitable studies was challenging due to the age range criteria. For this meta-analysis, it was essential to focus exclusively on neonates to study the effects of prenatal language exposure, as infants older than a few days may have already been exposed to their native language outside the womb.

2. Discrimination task between the native language and an unfamiliar language.

3. Stimuli include real speech (not single syllables or phonemes, speech analogues, synthesized speech or manipulated speech stimuli etc.).

During the literature search, studies using artificial languages or artificial grammar were also filtered out. Studies that explored the extent to which neonates discriminate against phonemes were also excluded. It was imperative that the stimuli of the included studies must have the native language and non-native or non-linguistic stimuli.

4. Monolingual infants (i.e. infants not exposed prenatally to both of the tested languages).

   Although there are bilingual studies that investigate the same topic, only monolingual studies were included in the meta-analysis. Recruiting monolingual newborns reduces the variability in language exposure and eliminates the confounding effects of exposure to multiple languages.

5. The use of NIRS (also known as OT, fNIRS) as the primary neuroimaging technique.

   Behavioral studies were excluded because they did not examine the cortical activity associated with language exposure. Integrating EEG and fNIRS data involves extensive analysis and effort to align the two modalities through standardizing the effect sizes and temporal and spatial characteristics. Due to these complications, EEG studies were also excluded from the meta-analysis.

Applying these selection criteria narrowed the studies down to eight. All eight studies were published. After reaching out to the authors, only the authors of three studies provided me with data. The other authors did provide data, because (i) they did not reply to my request, (ii) couldn’t share the data from their studies due to ethical constraints (no consent for data sharing) or (iii) difficulties locating the original datasets.

3.3 The Studies Selected and Participants

3.3.1 Studies

The three studies for which data was available were drawn from two articles authored by May et al. (2011) and May et al. (2017). The three studies had the same stimuli and experimental designs. May et al. (2010) compared the native language (English) to a non-native language (Tagalog) by using backwards speech as a control. In the first study of May et al. (2017), the native language (English) was compared to the non-native language (Spanish). In the second study of May et al. (2017), the non-native language (Spanish) was compared to a surrogate language (Silbo Gomero). Both studies in May et al. (2017) used backwards speech as a control.

In all the studies, the infants were tested using a NIRS device while they were sleeping or in a calm state. Testing occurred in a silent experimental room allowing the neonate to hear only the sound administered through two loudspeakers. "A Hitachi ETG-4000 NIRS machine with a source detector separation of 3 cm and two continuous wavelengths of 695 and 830 nm was used to record the NIRS signal, using a sampling rate of 10 Hz" was used for all three studies (May et al., 2011). Two chevron-shaped optical probes, each consisting of nine one mm optical fibers, were placed over the participant's head. One probe was placed over each hemisphere formed 12 optical channels for measurement. The relevant region of interest (ROI) was predefined by the initial studies, and if no ROI was found the mean of all channels was calculated.

3.3.2 Participants

The demographic information of the participants included in the meta-analysis are shown in Table 1. A total of 64 infants were included.

Table 1

Participant Demographics

Study	**Sample size**	**Native Language**	**Age Range**	**Mean Age**
May et al., 2011	20	English	0-3 days	1.6 days
May et al., 2017 Study 1	24	English	0-3 days	1.46 days
May et al., 2017 Study 2	20	English	0-4 days	1.21 days
Total	**64**	**English**	**0-4 days**	**1.23 days**

Note: Sample size represents the number of neonates in each study. Mean ages are calculated in days.

3.4 Data Analysis

3.4.1 fNIRS Preprocessing

All studies were preprocessed in the same way as in the original publications, as I used their preprocessed data. Specifically, in May et al. (2011), the analyses were conducted between 0 to 35 seconds after the stimulus onset in order to capture the full-time course of the hemodynamic response in each block. The data was band-pass filtered between 0.01 and 0.7 Hz to remove low-frequency and high-frequency noise. By isolating blocks where a change in concentration was greater than 0.1 mmol x mm over a period of .2 seconds, movement artifacts were detected and removed. On average, 3.46 blocks were retained for non-native language forward, 3.16 in non-native language backward, 3.69 for native language forward, and 3.17 for native language backward (May et al., 2011). A baseline was established for all the retained blocks.

Both studies in May et al. (2017) used the same procedure and analysis. Due to previous research, the authors focused on oxygenated hemoglobin as it is the strongest marker of neural activity in infant NIRS (Aslin, 2013; Gervain et al., 2011; Lloy-Fox et al., 2010). Between 4.2 to 17.1 seconds, the changes in oxygenated hemoglobin were examined after the initial stimulus onset. Similar to May et al. (2011) the data was filtered through 0.001 and 0.7 Hz to remove low-frequency and high-frequency noise. Movement artifacts were removed the same way as in May et al. (2011). A baseline was established for each trial.

3.4.2 Calculation of Effect Sizes

To calculate the effect size, I used the code developed by Gemignani et al. (2023) for conducting meta-analyses of fNIRS studies in MATLAB. The code supports two types of analyses: a mixed-effects modeling approach, which analyzes infant-level effect sizes, and a meta-analytical approach, which focuses on study-level effect sizes. For this thesis, only the study-level meta-analytical approach was applied (Gemignani et al., 2023).

Individual means were calculated by averaging each participant’s response across all trials for a given condition. These means were then standardized by dividing them by the standard deviation, creating normalized effect size estimates that account for the differences in response variability across participants and studies.

The sampling variances of the meta-analytic effect sizes were calculated by using the formula:

$$V_d = \frac{2}{n} + \frac{d_{\text{study}}^2}{4n},$$

with *d* representing the meta-analytic effect sizes and *n* representing the number of participants in each study (Gemignani et al., 2023). This formula accounts for both the number of participants and the magnitude of the effect size, providing a weighted average that reflects the precision of the effect size for each study (Gemignani et al., 2023).

After the code computed the effect sizes for each channel and hemoglobin component independently, I averaged the effect sizes across channels within each study's defined region of interest (ROI). If a study did not specify a ROI, the average effect size across all channels was computed to ensure a comprehensive analysis. Overall, this approach allowed for a robust analysis of the fNIRS data across the three studies, providing insights into the cortical responses to the different conditions.

3.4.3 Statistical Analysis

The study-level effect sizes were analyzed using the classical meta-analysis module in JASP. JASP automatically computed the overall effect size, confidence intervals, and residual heterogeneity estimates.

Four separate meta-analyses were conducted, one for each of the following comparisons:

(i) the comparison between the native language to backwards speech

(ii) the comparison between the non-native language to backwards speech

(iii) the comparison between the native language to the non-native language

(iv) the comparison between the backwards speech of the native language to the backwards speech of the non-native language

The second study in May et al. (2017) was only included in the comparison between the non-native language and backwards directed speech, as it did not include the native language of the participants. The first study in May et al. (2017) and May et al. (2011) were included in all comparisons as they investigated both the native language and non-native language of the infants.

Hemisphere was included as a factor in each of the four analyses. The hypothesis of this thesis was that the left hemisphere should show significant activation, indicating a pattern of left hemisphere involvement in language processing similar to that observed in adults. If the left hemisphere has significance only in the native language and not the non-native language, then prenatal language exposure does shape the neural network of the developing fetus. Both the oxygenated hemoglobin and deoxygenated hemoglobin variables were tested, despite oxygenated hemoglobin being the strongest marker of neural activity in infant NIRS studies (Aslin, 2013; Gervain et al., 2011). Both oxygenated and deoxygenated hemoglobin variables were analyzed to determine if there were any significant effects present.

4 Results

4.1 Oxygenated Hemoglobin (HbO)

4.1.1 Forward (FW) Native Language v Backward (BW) Native Language

The results of the first meta-analysis comparing FW and BW native language are shown in Figure 1 below. The mean effect size across the two studies is 0.26, which is significant ($p < 0.001$), with no explained residual variance.

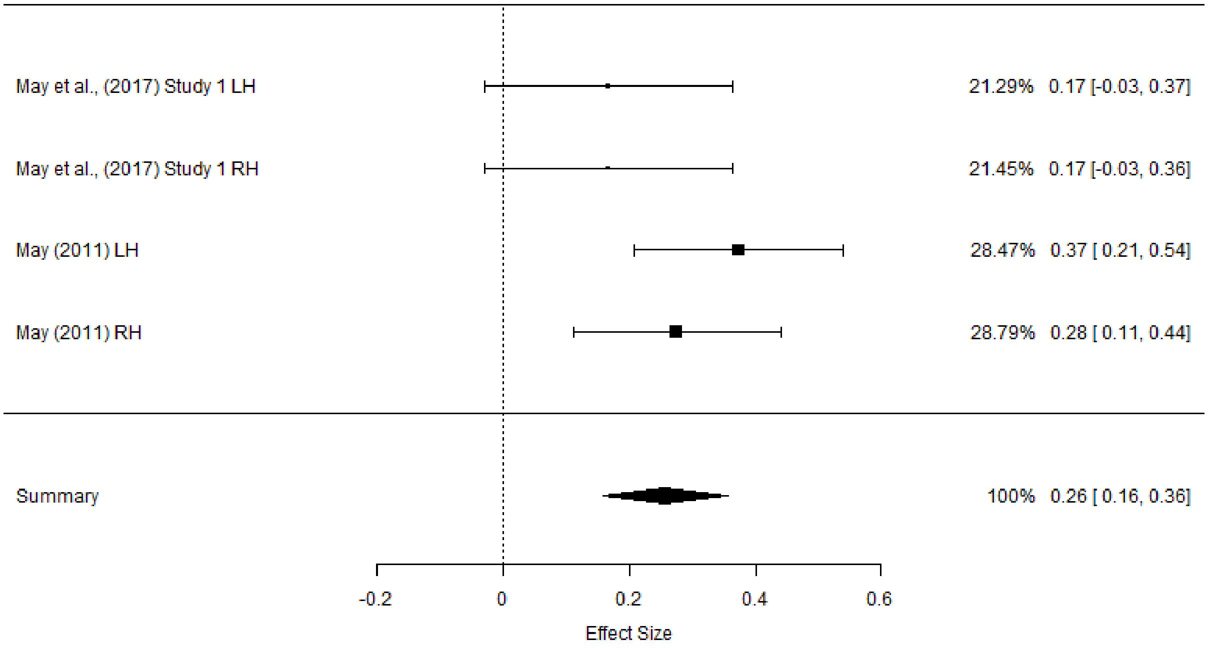

Figure 1. Forest plot for the FW vs. BW native language meta-analysis.

A Rosenthal's fail-safe N of 41 means that it would take 41 unpublished or missing studies with null results to reduce the overall effect size to a non-significant value. Due to how high this number is, it is highly unlikely that it will be overturned by missing studies.

4.1.2 FW Non-Native Language v BW Non-Native Language

The results of the second meta-analysis comparing FW and BW of the non-native language are shown in Figure 2 below. The mean effect size across the three studies is -0.08, which is insignificant (p = 0.099). There is a significant unexplained variability in the effect size across studies.

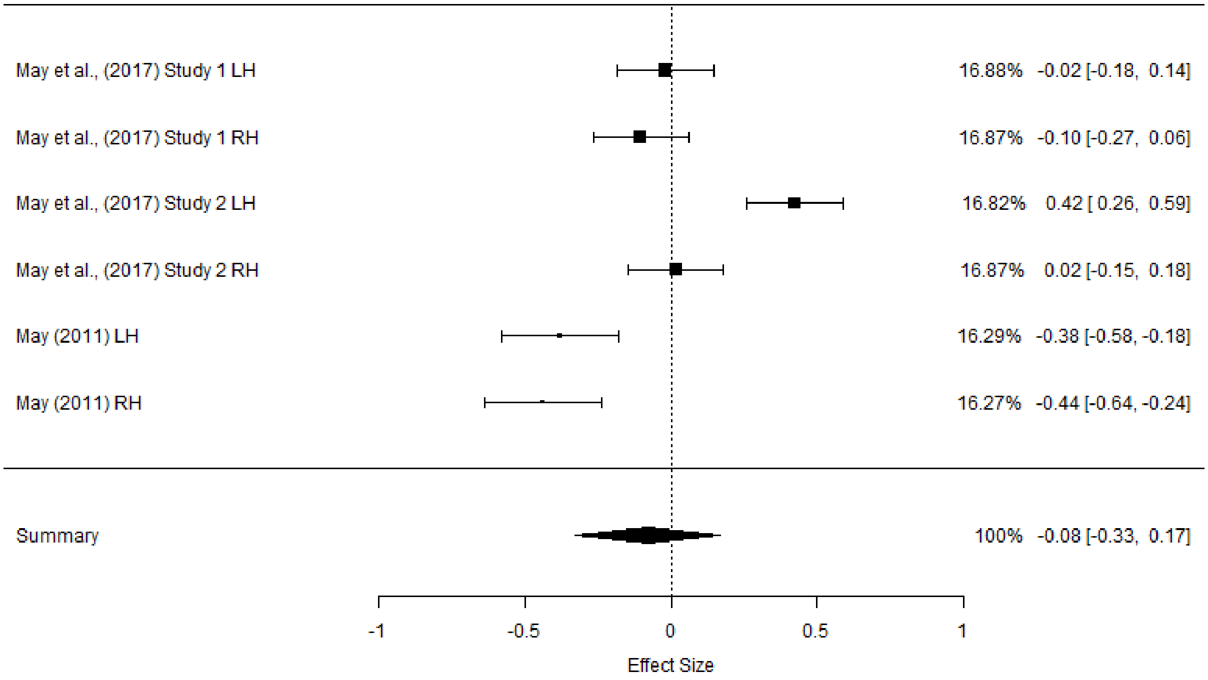

Figure 2. Forest plot for the FW vs. BW non-native language meta-analysis.

4.1.3 FW Native Language v FW Non-Native Language

The results of the third meta-analysis comparing FW native language and FW non-native language are shown in Figure 3 below. The mean effect size across the two studies is 0.30, which is significant ($p < 0.001$), with no explained residual variance.

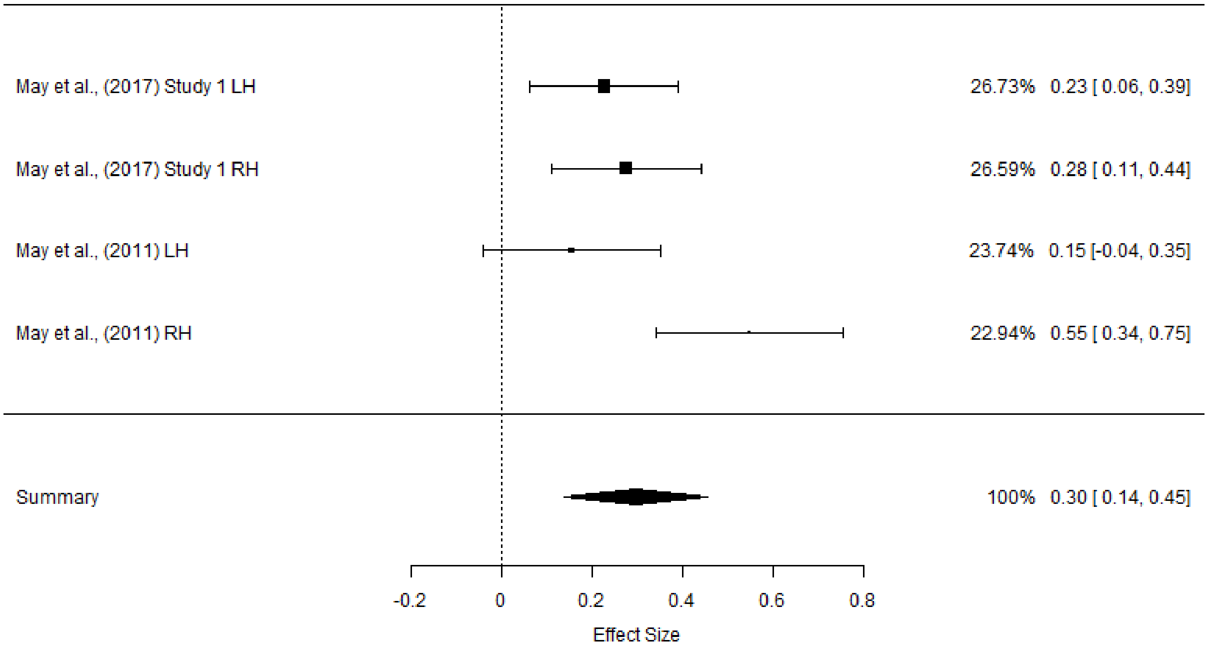

Figure 3. Forest plot for the FW native vs. FW non-native language meta-analysis.

A fail-safe N of 56 is extremely high, indicating that the results are very robust. The observed significance being lower than .001 indicates that the overall results are highly significant and that the effect size is unlikely to be overturned but missing or unpublished studies.

4.1.4 BW Native Language v BW Non-Native Language

The results of the fourth meta-analysis comparing BW native language and BW non-native language are shown in Figure 4 below. The mean effect size across the two studies is -0.24, which is significant ($p < 0.001$), with no explained residual variance.

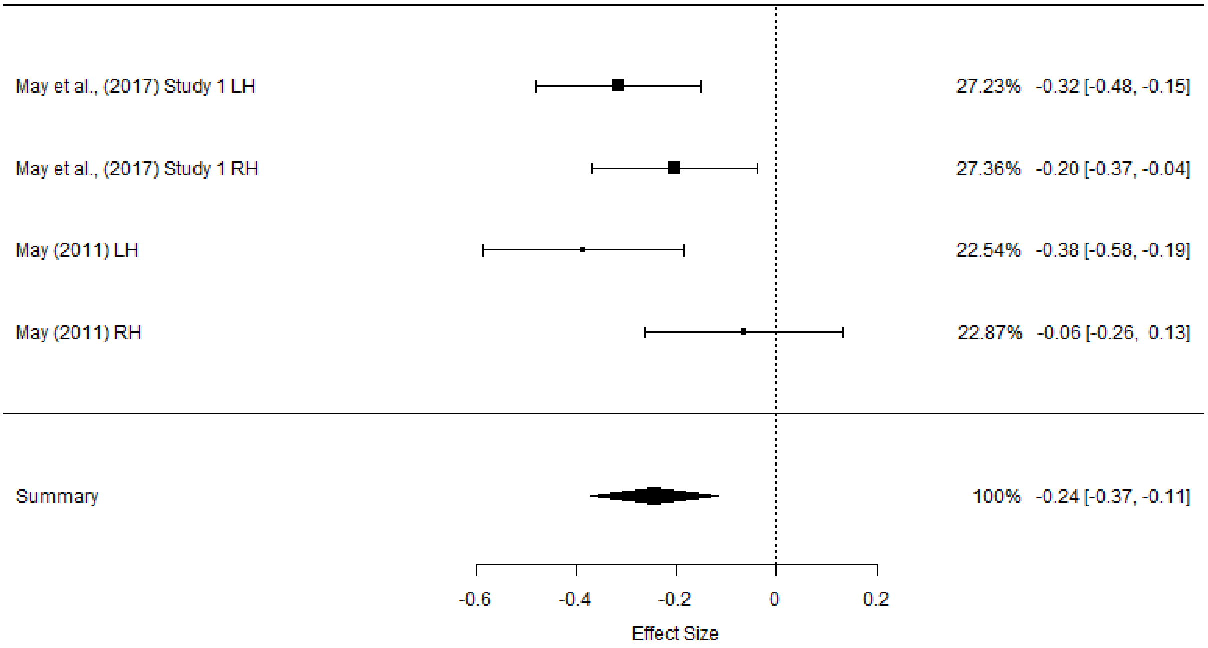

Figure 4. Forest plot for the BW native vs. BW non-native language meta-analysis.

The high fail-safe N shows that the results are very robust and unlikely to be influenced by unpublished studies. With fail-0safe N being 38, that means there needs to be an estimate of 38 null studies to reduce the observed effect to become non-significant. The observed significance is less than .001, which is well below the target significance of 0.050. This indicates that the results are highly significant, and the overall findings are strong and reliable.

In conclusion, these results suggest that both hemispheres contribute significantly, with the left hemisphere showing a stronger effect, and the model is reliable with no unexplained heterogeneity between the studies.

4.2 Deoxygenated Hemoglobin (HbR)

4.2.1 FW Native Language v BW Native Language

The results of the first meta-analysis comparing FW native language and BW native language with the HbR data are shown in Figure 5 below. The mean effect size across the two studies is -0.09, which is significant (p = 0.004). There are some residual variances, suggesting that the effect sizes vary more than would be expected by chance alone (p = 0.002)

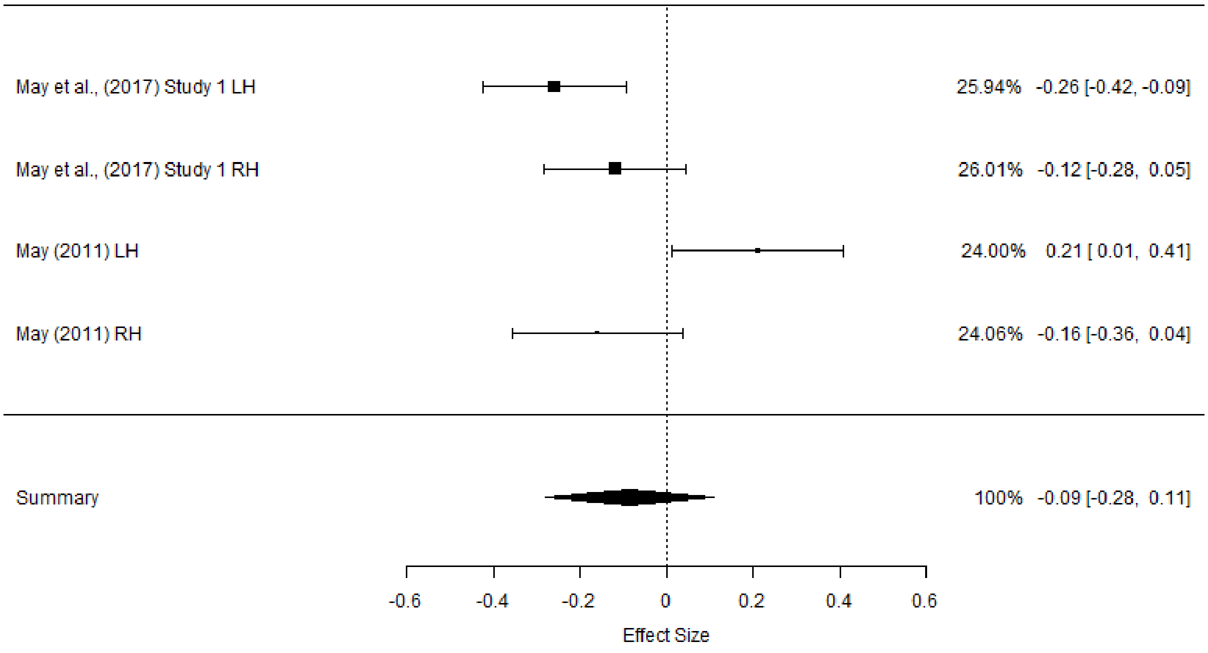

Figure 5. Forest plot for the FW vs. BW native language meta-analysis.

This forest plot shows that the left hemispheres in both studies show a significant effect but are in opposite directions. The first study in May et al. (2017) being negative and May et al. (2011) being positive. The right hemispheres in both studies are not significant as they are close to zero.

4.2.2 FW Non-Native Language v BW Non-Native Language

The results of the second meta-analysis comparing FW and BW of the non-native language with the HbR data are shown in Figure 6 below. The mean effect size across the three studies is 0.14, which is significant (p < 0.001). There is a significant unexplained variability in the effect size across studies, indicating there may be additional factors influencing the effect sizes across studies.

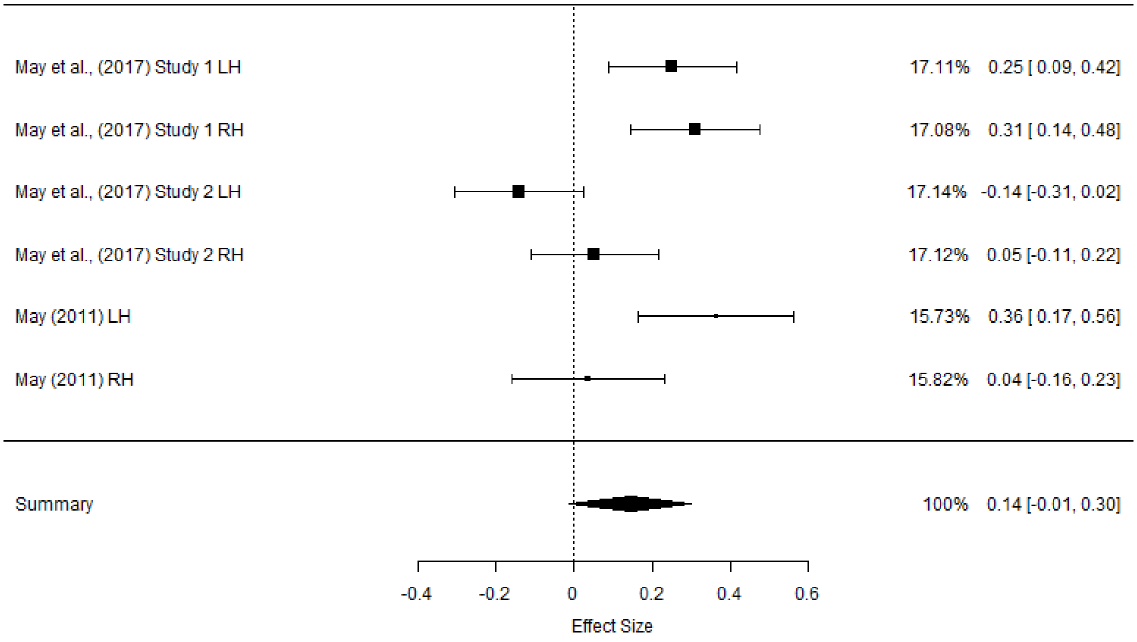

Figure 6. Forest plot for the FW vs. BW non-native language meta-analysis.

In this forest plot, only the left hemisphere is significant. The left hemisphere has a positive effect in the first study of May et al. (2017) and May et al. (2011), but a negative effect in the second study of May et al. (2017).

4.2.3 FW Native Language v FW Non-Native Language

The results of the third meta-analysis comparing FW native language and FW non-native language with the HbR data are shown in Figure 7 below. The mean effect size across the two studies is -0.18., which is significant ($p < 0.001$), with minimal unexplained residual variance.

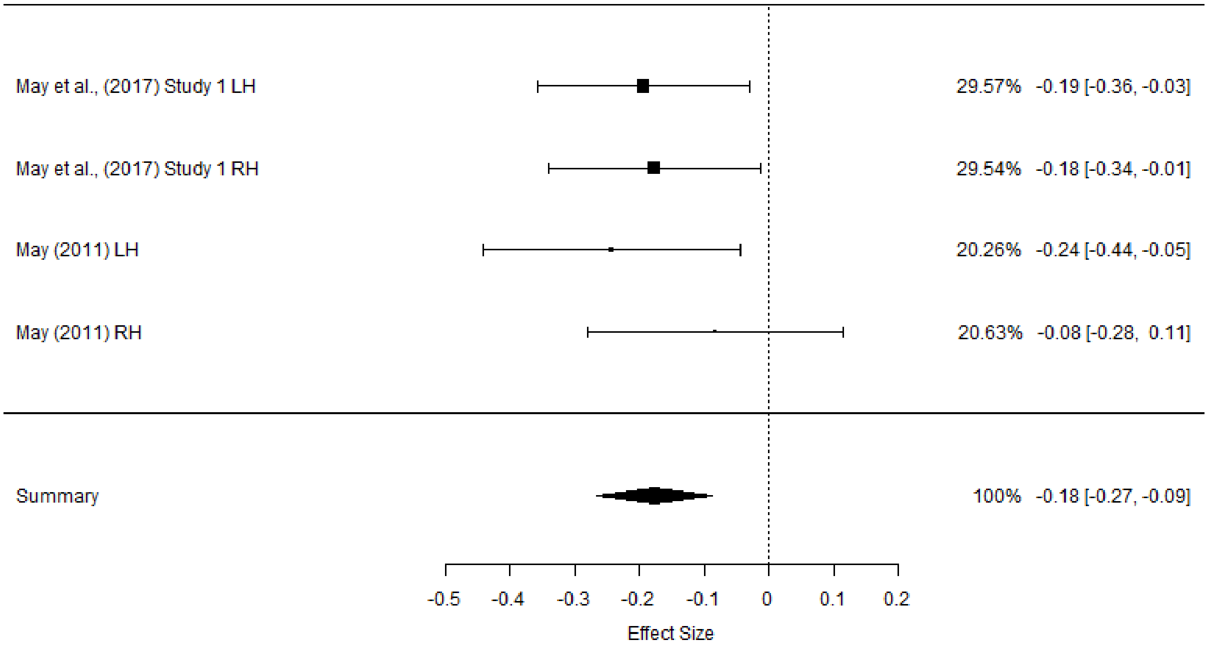

Figure 7. Forest plot for the FW native vs. FW non-native language meta-analysis.

The forest plot confirms that both hemispheres show significant negative effects in the first study in May et al. (2017), while only the left hemisphere in May et al. (2011) shows a significant negative effect.

4.2.4 BW Native Language v BW Non-Native Language

The results of the fourth meta-analysis comparing BW native language and BW non-native language with the HbR data are shown in Figure 8 below. The mean effect size across the two studies is 0.18, which is significant ($p < 0.001$), with significant unexplained residual variance.

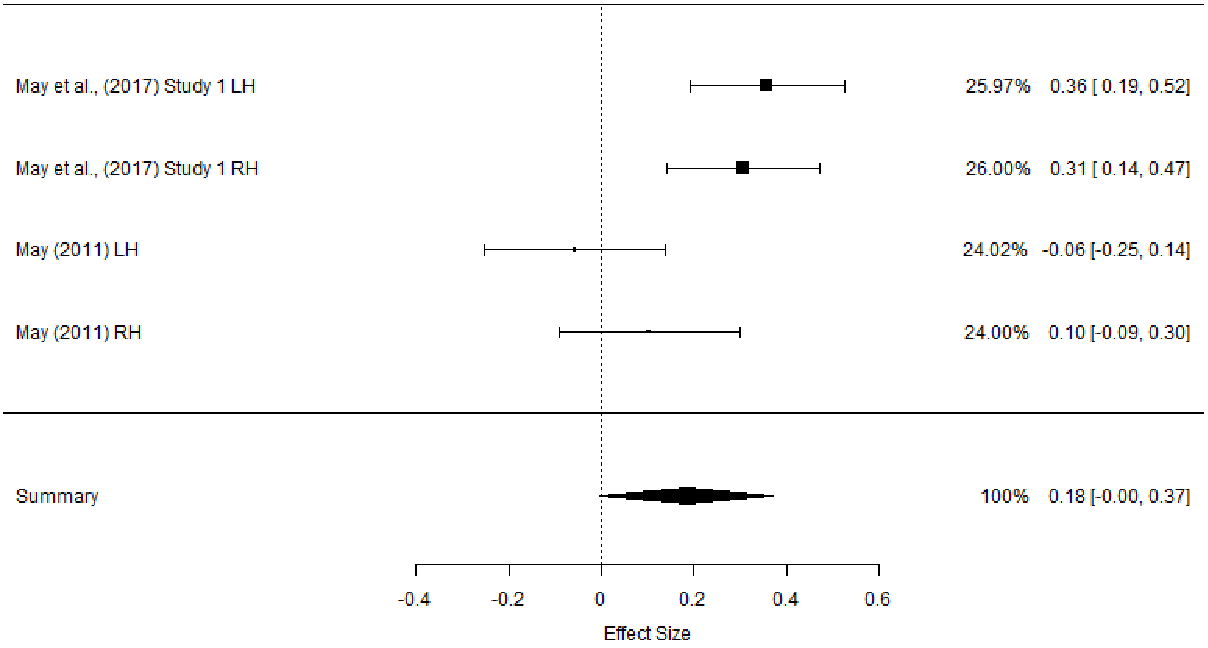

Figure 8. Forest plot for the BW native vs. BW non-native language meta-analysis.

In this graph only the first study in May et al. (2017) shows significant effect in both hemispheres, but this is not the case for May et al. (2011). This means that only one study showed significant hemispheric effects.

5 Discussion

The primary aim of this thesis is to synthesize data from studies investigating the effects of prenatal language exposure on the development of the auditory cortex in neonates. By utilizing fNIRS, these studies compared the neural responses of neonates to their native language (the language they were exposed to in utero), non-native language, and backward speech in both the native and non-native languages. In adults, language processing predominantly occurs in the left superior temporal lobe and left inferior frontal, and in the neural pathways connecting them, indicating left-hemisphere lateralization for language. By examining neonates, researchers can determine whether their brains process all languages in the left hemisphere, or if this lateralization is specific to the language they were exposed to prenatally. Given the pivotal role of the left hemisphere in language processing, this meta-analysis focused on comparing left and right hemisphere activations, and testing whether it is specifically the prenatally heard language that is processed preferentially by the left hemisphere, and whether an advantage for forward as compared to backward speech exists for the native language.

After analyzing the results of the four speech comparisons, any comparison that included the native language held significance. These consistent results were observed across the three studies. This occurred in the forward versus backwards speech for native speech, the forward native speech versus forward non-native speech, and the backward native speech versus backward non-native speech. The native language consistently activated both the left hemisphere and right hemisphere, with significantly stronger activation in the left hemisphere, suggesting left lateralization for the native language. These findings provide strong support for the prosodic bootstrapping theory, suggesting that prenatal language exposure has a direct impact on the prenatal development of the auditory cortex.

In the forest plot analysis of all languages, Tagalog in May et al. (2011) showed a negative effect size in both hemispheres, whereas Spanish in the second study in May et al. (2017) had a significantly positive effect size in the left hemisphere. This positive effect was similar with the effects observed for the native language. May et al. (2017) hypothesized that this could be attributed to context. In the first study neonates were exposed to their native language, eliciting a left hemisphere activation, while in the second study they were exposed to a surrogate language. Silbo Gomero was not processed as a language, similar to how Spanish (first study in May et al., 2017) and Tagalog (May et al., 2011) were processed. Since Spanish was the only language stimulus in the second study, it was processed as a language by the left hemisphere. Further research would be required to confirm whether this phenomenon holds across other languages and contexts.

When comparing the native and non-native forward speech, significant activation was observed in the right hemisphere. This may suggest that the right hemisphere responded similarly in both conditions, whereas the left hemisphere showed differential activation depending on the condition. This may be due to the left hemisphere exhibiting greater activation during the native language condition compared to the non-native language condition. The significant negative effect observed when comparing native and non-native backward-directed speech further supports the effectiveness of backward speech as a control, as it was processed consistently across all studies.

Research literature suggests that deoxygenated hemoglobin (HbR) typically shows weaker effects in infants than HbO. In the HbR results, significant heterogeneity was found across the comparison of forward and backward speech in both native and non-native languages, as well as in the backward comparisons between the two languages. Only in the comparison between forward native and forward non-native language was there little to no heterogeneity found. Overall, the HbR results were insignificant compared to the more reliable HbO data.

5.1 Limitations and Future Studies

The main limitation of this meta-analysis is that it includes only three studies, all published by the same authors. The small sample size limits the generalizability of the findings, especially since the native language in all studies was English. Other studies that met the inclusion criteria but were not analyzed involved neonates with different native languages such as Italian, Finnish, Japanese, and French (Pena et al., 2003; Kotilahti et al., 2010; Sato et al., 2011; Vannasing et al., 2016). Since these languages belong to different rhythmic classes, comparing them to English, a stress-timed language, could yield deeper insights into prenatal prosodic information exposure and rhythmic classes in early language acquisition.

The limited number of studies reduces statistical power and increases the potential for publication bias, as each study contributes more weight to the overall effect size. With only three studies, the risk of heterogeneity increases, making it challenging to isolate the true effect size. Additionally, there is a potential risk of over-interpretating findings due to the limited dataset. Although eight studies met the inclusion criteria, only three provided raw data, introducing the potential for file drawer effects, where unpublished or inaccessible studies with null results could impact the analysis if available. For future studies, it will be essential to incorporate all available published and unpublished studies that meet the inclusion criteria.

6 Conclusion

The adult brain exhibits left-hemisphere lateralization while processing familiar languages. Does this lateralization occur gradually over the years as we are exposed to our native language(s) or are our brains left lateralized to all languages at birth and we lose this ability as we are exposed to our native language. The auditory system becomes functional while the fetus is still developing inside the womb. Low pass filtered sounds, including speech, can flow through the maternal tissue and amniotic fluid to reach the fetus. This means that the fetus is already exposed to the prosodic information of their mother tongue even before they are born. The aim of this meta-analysis was to investigate prenatal language acquisition by synthesizing the findings of neonatal fNIRS studies.

The studies used in this meta-analysis tested the neonates using their native language and a language belonging to a different rhythmic class. In these studies, backwards edited speech of either the native or non-native language was used as a control. In the meta-analysis, backward-directed speech was compared with both native and non-native languages, as well as between native and non-native backward-directed speech.

The results of this meta-analysis varied slightly from the original studies; however, they were largely consistent overall. The native language (English) was consistently significant whether compared to backwards speech, the non-native language, or a surrogate language (Silbo). Compared to its backwards speech, the native language showed significant activation in the left and right hemisphere, more so in the left. The non-native language (Spanish and Tagalog) showed no significance in either hemisphere, when compared to its backwards speech counterpart. Backward native versus backward non-native showed significance in both the left and right hemisphere. This would suggest that both hemispheres were active during this comparison.

In conclusion, this meta-analysis compared the results of three studies on neonates using fNIRS. The findings support the prosodic bootstrapping theory of language acquisition, as the native language elicited greater left hemisphere activation compared to the non-native language. This suggests that the language the neonate was exposed to in utero impacts the development of the auditory cortex and language processing centers in the brain.

References

Abrams, R. M., & Gerhardt, K. J. (1996). Vibration of the abdominal segment in pregnant sheep. *Seminars in perinatology*, *20*(1), 30–37. https://doi.org/10.1016/s0146-0005(96)80055-3

Arimitsu, T., Minagawa, Y., Yagihashi, T., Uchida, M. O., Matsuzaki, A., Ikeda, K., & Takahashi, T. (2018). The cerebral hemodynamic response to phonetic changes of speech in preterm and term infants: The impact of postmenstrual age. NeuroImage: Clinical, 19, 599-606. https://doi.org/10.1016/j.nicl.2018.05.005

Aslin R. N. (2012). Questioning the questions that have been asked about the infant brain using near-infrared spectroscopy. *Cognitive neuropsychology*, *29*(1-2), 7–33. https://doi.org/10.1080/02643294.2012.654773

Aslin, R. N., Shukla, M., & Emberson, L. L. (2015). Hemodynamic correlates of cognition in human infants. *Annual review of psychology*, *66*, 349–379. https://doi.org/10.1146/annurev-psych-010213-115108

Baker, W. B., Parthasarathy, A. B., Busch, D. R., Mesquita, R. C., Greenberg, J. H., & Yodh, A. G. (2014). Modified Beer-Lambert law for blood flow. *Biomedical Optics Express, 5*(11), 4053-4075. https://doi.org/10.1364/BOE.5.004053

Bartha-Doering, L., Alexopoulos, J., Giordano, V., Stelzer, L., Kainz, T., Benavides-Varela, S., Wartenburger, I., Klebermass-Schrehof, K., Olischar, M., Seidl, R., & Berger, A. (2019). Absence of neural speech discrimination in preterm infants at term-equivalent age. *Developmental cognitive neuroscience*, *39*, 100679. https://doi.org/10.1016/j.dcn.2019.100679

Butler, J., Floccia, C., Goslin, J., & Panneton, R. (2011). Infants' discrimination of familiar and unfamiliar accents in speech. Infancy, 16(4), 392–417.

Carreiras, M., Lopez, J., Rivero, F., & Corina, D. (2005). Linguistic perception: Neural processing of a whistled language. Nature, 433.

Chen, X., Affourtit, J., Ryskin, R., Regev, T. I., Norman-Haignere, S., Jouravlev, O., Malik-Moraleda, S., Kean, H., Varley, R., & Fedorenko, E. (2023). The human language system, including its inferior frontal component in "Broca's area," does not support music perception. *Cerebral Cortex, 33*(12), 7904–7929. https://doi.org/10.1093/cercor/bhad087

Choi D, Black AK, Werker JF. (2018). Cascading and multisensory influences on speech perception development. Mind Brain Educ. 12(4):212–23

Chomsky, N. (1959). A review of B. F. Skinner's Verbal Behavior. Language, 35(1), 26-58.

Chomsky, N. (1965). Aspects of the theory of syntax. MIT Press.

Chomsky, N. (1981). Lectures on Government and Binding: The Pisa Lectures. Mouton de Gruyter.

Chomsky, N. (1986). Knowledge of Language: Its Nature, Origin, and Use. Praeger.

DeCasper, A. J., & Fifer, W. P. (1980). Of human bonding: Newborns prefer their mothers' voices. *Science, 208*(4448), 1174–1176. https://doi.org/10.1126/science.7375928

Dehaene-Lambertz, G., Dehaene, S., & Hertz-Pannier, L. (2002). Functional neuroimaging of speech perception in infants. Science (New York, N.Y.), 298(5600), 2013–2015. https://doi.org/10.1126/science.1077066

Dorn, K., Cauvet, E., & Weinert, S. (2019). A cross-linguistic study of multisensory perceptual narrowing in German and Swedish infants during the first year of life. *Infant and Child Development, 30*(1), e2217. https://doi.org/10.1002/icd.2217

Eggermont JJ, Moore JK. 2012. Morphological and functional development of the auditory nervous system. In Human Auditory Development, ed. L Werner, RR Fay, AN Popper, pp. 61–105. New York: Springer.

Ekström, A. G. (2024). Correcting the record: Phonetic potential of primate vocal tracts and the legacy of Philip Lieberman (1934−2022). *American Journal of Primatology*, 86(8), e23637. https://doi.org/10.1002/ajp.23637

Gasparini, L., Langus, A., Tsuji, S., & Boll-Avetisyan, N. (2021). Quantifying the role of rhythm in infants' language discrimination abilities: A meta-analysis. Cognition, 213, 104757. https://doi.org/10.1016/j.cognition.2021.104757

Gaynutdinova, A., & Mutallimova, A. (2021). The Culture and Language of Whistle of Turkish People (Giresun). *International Journal of Society, Culture & Language*, *9*(1), 86-93.

Gemignani, J., & Gervain, J. (2021). Comparing different pre-processing routines for infant fNIRS data. *Developmental Cognitive Neuroscience, 48,* 100943. https://doi.org/10.1016/j.dcn.2021.100943

Gemignani, J., de la Cruz-Pavía, I., Martinez, A., Nallet, C., Pasquini, A., Lucarini, G., Cavicchiolo, F., & Gervain, J. (2023). Reproducibility of infant fNIRS studies: A meta-analytic approach. *Neurophotonics, 10*(2), 1–21. https://doi.org/10.1117/1.NPh.10.2.023518

Gerhardt, K. J., & Abrams, R. M. (2000). Fetal exposures to sound and vibroacoustic stimulation. Journal of perinatology : official journal of the California Perinatal Association, 20(8 Pt 2), S21–S30. https://doi.org/10.1038/sj.jp.7200446

Gervain, J., Nespor, M., Mazuka, R., Horie, R., & Mehler, J. (2008). Bootstrapping word order in prelexical infants: a Japanese-Italian cross-linguistic study. *Cognitive psychology*, *57*(1), 56–74. https://doi.org/10.1016/j.cogpsych.2007.12.001

Gervain J, Mehler J. (2010). Speech perception and language acquisition in the first year of life. Annu. Rev. Psychol. 61:191–218

Gervain, J., Mehler, J., Werker, J. F., Nelson, C. A., Csibra, G., Lloyd-Fox, S., Shukla, M., & Aslin, R. N. (2011). Near-infrared spectroscopy: A report from the McDonnell infant methodology consortium. *Developmental Cognitive Neuroscience, 1*(1), 22–46. https://doi.org/10.1016/j.dcn.2010.07.004

Gervain, J. (2018). The role of prenatal experience in language development. *Current Opinion in Behavioral Sciences, 21*, 62-67. https://doi.org/10.1016/j.cobeha.2018.02.004

González, P. V., Güntürkün, O., & Ocklenburg, S. (2020). Lateralization of auditory processing of Silbo Gomero. *Symmetry, 12*(7), 1183. https://doi.org/10.3390/sym12071183

Grabe, E. & Low, E. (2002). Durational variability in speech and the Rhythm Class Hypothesis. In C. Gussenhoven & N. Warner (Ed.), *Laboratory Phonology 7* (pp. 515-546). Berlin, New York: De Gruyter Mouton. https://doi.org/10.1515/9783110197105.2.515

Hepper, P. G., & Shahidullah, B. S. (1994). Development of fetal hearing. Archives of disease in childhood. Fetal and neonatal edition, 71(2), F81–F87. https://doi.org/10.1136/fn.71.2.f81

Hohle, B., Bijeljac-Babic, R., & Nazzi, T. (2020). Variability and stability in early language acquisition: Comparing monolingual and bilingual infants' speech perception and word recognition. Bilingualism: Language and Cognition, 23(1), 56–71.

James, L. (2021). Systems of communication: Aspects of culture and structure in speech surrogates. *Frontiers in Communication, 6*, 653268. https://doi.org/10.3389/fcomm.2021.653268

Jeong, H., Sugiura, M., Sassa, Y., Haji, T., Usui, N., Taira, M., & Kawashima, R. (2007). Cross-linguistic influence on brain activation during second language processing: An fMRI study. *Bilingualism: Language and Cognition, 10*(2), 175–187. https://doi.org/10.1017/S1366728907002921

Jungewelter, S., Taskinen, H., Sallmén, M., Lindbohm, M. L., Airo, E., Remes, J., Huotilainen, M., & Jansson-Verkasalo, E. (2024). Maternal occupational noise exposure during pregnancy and children's early language acquisition. PloS one, 19(4), e0301144. https://doi.org/10.1371/journal.pone.0301144

Kisilevsky, B. S., Hains, S. M., Brown, C. A., Lee, C. T., Cowperthwaite, B., Stutzman, S. S., Swansburg, M. L., Lee, K., Xie, X., Huang, H., Ye, H. H., Zhang, K., & Wang, Z. (2009).

Fetal sensitivity to properties of maternal speech and language. *Infant behavior & development*, *32*(1), 59–71. https://doi.org/10.1016/j.infbeh.2008.10.002

Kotilahti, K., Nissilä, I., Näsi, T., Lipiäinen, L., Noponen, T., Meriläinen, P., Huotilainen, M., & Fellman, V. (2010). Hemodynamic responses to speech and music in newborn infants. *Human brain mapping*, *31*(4), 595–603. https://doi.org/10.1002/hbm.20890

Kuhl, P. K. (2000). A new view of language acquisition. *Proceedings of the National Academy of Sciences, 97*(22), 11850-11857. https://doi.org/10.1073/pnas.97.22.11850

Lecanuet J-P, Granier-Deferre C. (1993). Speech stimuli in the fetal environment. In Developmental Neurocognition: Speech and Face Processing in the First Year of Life, ed. B Boysson-Bardies, S de Schonen, P Jusczyk, P MacNeilage, J Morton, pp. 237–48. Dordrecht, Neth.: Springer

Li, R., Yang, D., Fang, F., Hong, K.-S., Reiss, A. L., & Zhang, Y. (2022). Concurrent fNIRS and EEG for brain function investigation: A systematic, methodology-focused review. *Sensors, 22*(15), 5865. https://doi.org/10.3390/s22155865

May, L., Byers-Heinlein, K., Gervain, J., & Werker, J. F. (2011). Language and the newborn brain: does prenatal language experience shape the neonate neural response to speech?. *Frontiers in psychology*, *2*, 222. https://doi.org/10.3389/fpsyg.2011.00222

May, L., Gervain, J., Carreiras, M., & Werker, J. F. (2017). The specificity of the neural response to speech at birth. *Developmental science*, *21*(3), e12564. https://doi.org/10.1111/desc.12564

Minagawa-Kawai, Y., van der Lely, H., Ramus, F., Sato, Y., Mazuka, R., & Dupoux, E. (2011). Optical brain imaging reveals general auditory and language-specific processing in early infant development. Cerebral cortex (New York, N.Y. : 1991), 21(2), 254–261. https://doi.org/10.1093/cercor/bhq082

Mercure, E., Evans, S., Pirazzoli, L., Goldberg, L., Bowden-Howl, H., Coulson-Thaker, K., Beedie, I., Lloyd-Fox, S., Johnson, M. H., & MacSweeney, M. (2020). Language Experience Impacts Brain Activation for Spoken and Signed Language in Infancy: Insights From Unimodal and Bimodal Bilinguals. *Neurobiology of language (Cambridge, Mass.)*, *1*(1), 9–32. https://doi.org/10.1162/nol_a_00001

Mushtaq, F., Wiggins, I. M., Kitterick, P. T., Anderson, C. A., & Hartley, D. E. H. (2019). Evaluating time-reversed speech and signal-correlated noise as auditory baselines for isolating speech-specific processing using fNIRS. *PLoS ONE, 14*(7), e0219927. https://doi.org/10.1371/journal.pone.0219927

Nallet, C., & Gervain, J. (2021). Neurodevelopmental preparedness for language in the neonatal brain. *Annual Review of Developmental Psychology, 3*, 41–58. https://doi.org/10.1146/annurev-devpsych-050620-025732

Nazzi, T., Bertoncini, J., & Mehler, J. (2000). Language discrimination by newborns: Toward an understanding of the role of rhythm. *Journal of Experimental Psychology: Human Perception and Performance, 26*(3), 996–1010. https://doi.org/10.1037/0096-1523.26.3.996

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

Nazzi, T., & Cutler, A. (2019). How consonants and vowels shape spoken-language recognition. Annual Review of Linguistics, 5(1), 25–47

Olulade, O. A., Seydell-Greenwald, A., Chambers, C. E., Turkeltaub, P. E., Dromerick, A. W., Berl, M. M., Gaillard, W. D., & Newport, E. L. (2020). The neural basis of language development: Changes in lateralization over age. *Proceedings of the National Academy of Sciences, 117*(38), 23477–23483. https://doi.org/10.1073/pnas.1905590117

Peña, M., Maki, A., Kovacic, D., Dehaene-Lambertz, G., Koizumi, H., Bouquet, F., & Mehler, J. (2003). Sounds and silence: An optical topography study of language recognition at birth. *Proceedings of the National Academy of Sciences, 100*(20), 11702-11705. https://doi.org/10.1073/pnas.1934290100

Peng, C., & Hou, X. (2021). Applications of functional near-infrared spectroscopy (fNIRS) in neonates. *Neuroscience Research, 170*, 18–23. https://doi.org/10.1016/j.neures.2020.11.003​:contentReference[oaicite:1]{index=1}

Pinti, P., Tachtsidis, I., Hamilton, A., Hirsch, J., Aichelburg, C., Gilbert, S., & Burgess, P. W. (2020). The present and future use of functional near-infrared spectroscopy (fNIRS) for cognitive neuroscience. *Annals of the New York Academy of Sciences*, *1464*(1), 5–29. https://doi.org/10.1111/nyas.13948

Ragó, A., Honbolygó, F., Róna, Z., Beke, A., & Csépe, V. (2014). Effect of maturation on suprasegmental speech processing in full- and preterm infants: A mismatch negativity study. *Research in Developmental Disabilities, 35*(1), 192-202. https://doi.org/10.1016/j.ridd.2013.10.006

Ramus, F., Nespor, M., & Mehler, J. (1999). Correlates of linguistic rhythm in the speech signal. *Cognition*, *73*(3), 265–292. https://doi.org/10.1016/s0010-0277(99)00058-x

Rimas, J., Rimas Jr, J. (2024). Music and Language (Connections Between Verbal and Musical Languages). In: Etudes on the Philosophy of Music. Palgrave Macmillan, Cham. https://doi.org/10.1007/978-3-031-63965-4_7

Saberi, K., & Perrott, D. R. (1999). Cognitive restoration of reversed speech. *Nature, 398*(6730), 760. https://doi.org/10.1038/19652

Sato, H., Hirabayashi, Y., Tsubokura, H., Kanai, M., Ashida, T., Konishi, I., Uchida-Ota, M., Konishi, Y., & Maki, A. (2011). Cerebral hemodynamics in newborn infants exposed to speech sounds: A whole-head optical topography study. *Human Brain Mapping, 32*(4), 594–604. https://doi.org/10.1002/hbm.21350

Skinner, B. F. (1957). *Verbal behavior*. Prentice Hall.

Skinner, B.F. (1969). Contingencies of reinforcement. A theoretical analysis. Appleton-Century-Crofts.

Silver, E., Korja, R., Mainela-Arnold, E., Pulli, E. P., Saukko, E., Nolvi, S., Kataja, E. L., Karlsson, L., Karlsson, H., & Tuulari, J. J. (2021). A systematic review of MRI studies of language development from birth to 2 years of age. *Developmental Neurobiology, 81*(1), 63–75. https://doi.org/10.1002/dneu.22792

Temperley, D. (2022). Music and language. *Annual Review of Linguistics, 8*(1), 153-170. https://doi.org/10.1146/annurev-linguistics-031220-121126

Tincoff, R., Hauser, M., Tsao, F., Spaepen, G., Ramus, F., & Mehler, J. (2005). The role of speech rhythm in language discrimination: Further tests with a non-human primate. *Developmental Science, 8*(1), 26-35. https://doi.org/10.1111/j.1467-7687.2005.00390.x

Ueda, K., Nakajima, Y., Ellermeier, W., & Kattner, F. (2017). Intelligibility of locally time-reversed speech: A multilingual comparison. *Scientific Reports, 7*, 1782. https://doi.org/10.1038/s41598-017-01831-z

Vannasing, P., Florea, O., González-Frankenberger, B., Tremblay, J., Paquette, N., Safi, D., Wallois, F., Lepore, F., Béland, R., Lassonde, M., & Gallagher, A. (2016). Distinct hemispheric specializations for native and non-native languages in one-day-old newborns identified by fNIRS. *Neuropsychologia, 84*, 63-69. https://doi.org/10.1016/j.neuropsychologia.2016.01.038

Vigneau, M., Beaucousin, V., Hervé, P. Y., Duffau, H., Crivello, F., Houdé, O., Mazoyer, B., & Tzourio-Mazoyer, N. (2006). Meta-analyzing left hemisphere language areas: phonology, semantics, and sentence processing. *NeuroImage*, *30*(4), 1414–1432. https://doi.org/10.1016/j.neuroimage.2005.11.002

Vogelsang, M., Vogelsang, L., Diamond, S., & Sinha, P. (2023). Prenatal auditory experience and its sequelae. Developmental science, 26(1), e13278. https://doi.org/10.1111/desc.13278

Vouloumanos, A., & Werker, J. F. (2004). Tuned to the signal: the privileged status of speech for young infants. *Developmental science*, *7*(3), 270–276. https://doi.org/10.1111/j.1467-7687.2004.00345.x

Vouloumanos, A., Hauser, M. D., Werker, J. F., & Martin, A. (2010). The tuning of human neonates' preference for speech. *Child development*, *81*(2), 517–527. https://doi.org/10.1111/j.1467-8624.2009.01412.x

Vuust, P., Heggli, O. A., Friston, K. J., & Kringelbach, M. L. (2022). Music in the brain. *Nature Reviews Neuroscience, 23*(5), 287-299. https://doi.org/10.1038/s41583-022-00578-5

Wilcox, T., & Biondi, M. (2015). fNIRS in the developmental sciences. *WIREs Cognitive Science, 6*(3), 263–283. https://doi.org/10.1002/wcs.1343