UNIVERSITÀ
DEGLI STUDI
DI PADOVA

Head Office: Università degli Studi di Padova

Department
Padova Neuroscience Center
Dipartimento di Psicologia dello Sviluppo e della Socializzazione

Ph.D. COURSE IN: NEUROSCIENCE – 36th cycle

THE SPECIAL STATUS OF LANGUAGE: NEURAL PROCESSING OF SPEECH IN NEWBORNS AND YOUNG INFANTS WITH TYPICAL AND ATYPICAL AUDITORY EXPERIENCE

Thesis written with the financial contribution of the ERC Consolidator Grant "BabyRhythm" no. 773202.

Coordinator: Prof. Antonino VALLESI
Supervisor: Prof. Judit GERVAIN
Co-Supervisor: Ch.mo Prof. Patrizia TREVISI

Ph.D. student : Caroline NALLET

The special status of language: Neural processing of speech in newborns and young infants with typical and atypical auditory experience

“Au matin vous entendez dans la rue des cris chantés,
qui vous indiquent l’heure qu’il est,
comme « du mouron pour les p’tits oiseaux »,
ou « harengs d’anuit », ou « merlans à frire, à frire », ou
« qui veut des toiles, voilà des toiles, voilà des toiles à laver ».
Si vous observez bien les sons qui vous arrivent,
vous constaterez que vous n’entendez pas des mots,
mais seulement une certaine chanson, que vous reconnaissez.
[…]
Le langage à distance est naturellement musical ;
il emprunte sa mélodie à l’accent du langage ordinaire ;
mais il la purifie et la simplifie afin d’être mieux compris,
et de plus loin.”

Alain, Propos.
26 octobre 1907.

[In the morning you hear the melodic cries in the street,
which tell you what time it is,
like “chickweeds for the li’l birds”,
or “night herrings” or “whitings to fry, to fry”, or
“who wants webs, here are some webs to wash”.
If you observe the sounds well,
you find that you don’t hear words,
but rather a kind of recognizable song.
[...]
Language from a distance is naturally musical;
It derives its melody from the accents of ordinary language;
but it’s made more pure and simple in order to be better understood,
and from further away.]

TABLE OF CONTENTS

List of figures and tables

Acknowledgments

First of all, I would like to thank Judit Gervain, the best thesis supervisor one could wish for. She introduced me to the fascinating world of infant research. Her ability to create a friendly, supportive and stimulating working environment allowed me to carry out this thesis work in the best possible conditions.

I would also like to thank all the members of my great research team: thank you to Caterina and Irene for having introduced me to research with babies during my internship in 2017 and who have continued to support me ever since, both in Padua or from Spain. Thank you to Marìa and Jessica and to Anna who was a great support during the first years of my PhD in Padova, marked by lockdowns but also by crispy sushis and funny cats. Thank you to Benedetta, Mohinish, Martina, Zeynep and Silvia: thanks for making this research team so great! A special thank to Gaia, my super co-worker on several projects, thank you for the wonderful work you did at the hospital and for all your sweetness.

Thank you to all the people who contributed to the achievement of my research work. Thanks to the doctors of the ENT service of the Hospital of Padova, Patrizia Trevisi, Davide Brotto and Alessandro Martini, for their availability and support. Thank you to all the medical team of the Necker's maternity in Paris and to prof. Alexandre Lapillonne which allows us to carry out exciting research with newborn babies and, of course, thank you to all the parents and babies who participated in my studies. Thank you also to the language research team of the INCC in Paris, where I started my "research career" in 2019. Finally, thanks to the persons who kindly hosted me during my stays in Paris: Sophie and Salomé, Cathy and Isa, Ana and Vasile, Merve, Louise.

Thank you to my friends: to the amazing Palace and to the kittens of Lyon. A special thank to Sarah that I am lucky to have by my side for many years now.

Thank you to my family. To my parents, for the love and pride I have always felt in their eyes, their words and their acts. To my big sisters and brother that have inspired me from the start; I feel very lucky to be the "little last one" of our siblings group. Thank you to my adorable nephews and niece, Satie, Tim and Pia and to Raphael and Louise. Thank you also to all the rest of my extended family, to Delphine and to Elie.

Thank you to Jeanne and Auguste for being so lovely. Their kindness and patience, their smartness and sense of humor have made it possible for me to carry out this work in such a gentle and joyful atmosphere.

Thank you to Jeannine Sombrero and Jeanne Charret for managing the Babylab of the Muguet so well and to Broutille for purring next to me while I was writing this thesis.

And thank you to Guillaume, my perfect lover…

1 Introduction

Newborns' speech perception abilities: what is speech at birth?

This chapter is based on the publication Nallet & Gervain (2021)

After only a year of exposure to speech, and without explicit training, infants start to produce their first words and soon afterwards short sentences, displaying a knowledge of language that increases considerably in only a few months. How do infants acquire their native language(s) so rapidly and easily? This question interests many researchers and numerous studies have highlighted the remarkable speech perception abilities that young infants display during their first months of life, before producing their first words, helping them discover their language from a very early stage.

The current thesis addresses one of the most fundamental and earliest tasks young learners need to face: How to identify the speech signal among the multitude of sounds they hear in their environment? What auditory stimuli do newborns categorize as speech and thus pay special attention to, in order to develop language? Does the newborn brain already exhibit a specific processing in response to speech compared to other auditory stimuli or some months of experience with language are required to shape the infant brain? The current thesis aims at bring new insights about these exciting questions. In order to address these questions, the thesis will first briefly review newborn infants' perceptual abilities (Chapter 1) and the experimental methods most commonly used to investigate them (Chapters 1 and 2). It will then present four empirical studies with newborns and young infants using the brain imaging technique near-infrared spectroscopy (NIRS) exploring how babies with typical and atypical hearing perceive speech and various other auditory stimuli. Experiment 1 (Chapter 3) will investigate how the newborn brain processes speech as well as newborn's cries. Experiment 2 (Chapter 3), as a subsequent study, will explore the newborn's brain processing of laughter. In experiment 3 (Chapter 4), we will study how the newborn's brain processes repetition-based regularities when presented with piano tones vs with syllables. Finally, experiment 4 (Chapter 5) will investigate how hearing loss influences the brain responses to languages, both native and unfamiliar.

1.1 Language acquisition : Why study newborns?

After the dominance of behaviorism, the 1950s–1960s brought about what was later termed the cognitive revolution—a new perspective on human psychology that emphasized internal mental representations and abilities even if they were not readily observable in overt behavior. Noam Chomsky's work on language was one of the foundations of this cognitive movement. Chomsky (1959) argued that humans use language productively, being able to express contents never heard before, and of theoretically infinite length (e.g., "Jim said that Ray believed that Robbie argued that John thought that. . . . "). This can be achieved only if children do not learn language by copying and imitating what is said in their environment, but rather possess a dedicated mental ability, i.e., a grammar, to generate language. Chomsky assumed that the ability to learn language was thus innate and that the young learner's task when acquiring a specific language (e.g., Chinese, French, or Zulu) was to configure this innate ability on the basis of the restricted input they hear. Chomsky drew a parallel between the language-learning child and the linguist who works on an unfamiliar language. They both need to establish the grammar and vocabulary of the target language on the basis of the linguistic data they observe. From a mathematical perspective, this task is unsolvable, since the learner runs into the induction problem: for any finite data set, there is an infinite number of underlying rule systems that could have generated them. It is, therefore, impossible to learn language on the basis of the input alone. If the child does not have innate guiding principles for choosing between the infinitely many possible rule systems, he or she will not be able to settle on any among the infinitely many options, and even if he or she chooses one rule system, there will be no guarantee that this will converge with those acquired by other speakers of the same language community. Yet, learners exposed to the same language in a given community develop functionally similar grammars. Chomsky concludes that given the finite nature of the input data set and the remarkable speed and ease with which human infants learn language, assuming an innate language faculty is a logical necessity.

This position was later countered by opposing empiricist theories (e.g., Elman et al. 1997, McCauley & Christiansen 2019, Tomasello 2000), which claimed that the input may be much richer than Chomsky argued, complete with not only large quantities of language input, but also convergent nonlinguistic cues, such as physical objects and, most importantly, social partners. As a consequence, they argued, language may be acquired through nonspecific learning and cognitive abilities, such as analogy making, statistical learning, theory of mind, and joint attention, without a need for assuming an innate language-specific ability.

As a consequence of this controversy (the nature-nurture debate), empirical research into young infants' speech perception, language learning, and other cognitive abilities has gained high theoretical relevance. Evidence for the early presence of an ability was often interpreted as evidence for its innate nature, whereas its absence early in development was taken to mean the ability is learned. Mapping the initial state of language development, before experience begins, has become particularly important. This has increased interest in investigating newborn infants' speech perception abilities, since newborns are the youngest population that can easily be tested in experimental studies. [Although methods exist to test fetuses, these options are more limited for obvious practical reasons (Dubois et al. 2014, Huotilainen et al. 2005, Weikum et al. 2012)].

Concurrently, attention has also been focused on what experience babies might have with language prenatally. Hearing is functional from about the 24th-28th week of gestation, depending on the specific auditory function measured (Eggermont & Moore 2012). Newborns thus already have experience with language before birth and do not represent the true initial state of language acquisition. Describing this experience is, therefore, important for interpreting the abilities observed in newborns. While directly observing the fetal auditory environment is challenging in humans, research with animal models, computational simulations, and fetal measures (e.g., Abrams & Gerhardt 2000, DeCasper & Sigafoos 1983, DeCasper et al. 1994, Gerhardt et al. 1992, Granier-Deferre et al. 2011, Griffiths et al. 1994, Querleu et al. 1988) suggests that maternal tissues and fluids act as low-pass filters. Fetuses thus receive a speech signal that is quite different from their postnatal speech experience and consists mainly of prosodic information. Prosody is the melody and rhythm of speech. How this prenatal experience impacts newborns' speech perception abilities is starting to be better understood (Mehler et al., 1988; Moon et al., 1993; Sato et al., 2012; Abboub et al., 2016).

With the advent of brain imaging techniques and advances in genetics and epigenetics, it has become increasingly clear that the strict dichotomy of nature or nurture cannot offer an appropriate framework for explaining language development. A more synergistic view has emerged emphasizing how genetically endowed and experience-dependent processes interact to bring about development (Choi et al. 2018, Gervain & Mehler 2010, Werker & Tees 2005). Under this view, the main focus is no longer to ask whether innate or learned mechanisms are responsible for language development, but rather to investigate how such mechanisms converge and interact.

In the current thesis, a better understanding of prenatal experience will help us investigate how newborns recognize speech: what are the acoustic properties that newborns use to identify an auditory stimulus as being speech, and thus a relevant stimulus to learn language? From birth, the newborn's experience with speech is only, or mostly, prosodic. How does this prosodic prenatal experience influence what newborns consider as speech before gaining experience with the full band speech signal?

1.2 How to study newborns?

Newborns, like young infants in general, are challenging research participants, with a limited attention span, a restricted behavioral repertoire, and an inability to follow instructions or to provide verbal answers. It is nevertheless possible to conduct behavioral experiments relying on newborns' spontaneous behaviors, and brain imaging techniques now allow us to probe the newborn brain noninvasively.

The most commonly used behavioral measures for the assessment of various perceptual and cognitive abilities in newborns include heart rate, a technique also used with fetuses (Della Longa et al. 2021, Nagy & Molnar 2004); looking times (Izard et al. 2009, Johnson et al. 1991); and high- amplitude (or nonnutritive) sucking (HAS) (Floccia et al. 1997). HAS, the method most frequently used to study speech perception in newborns, relies on a spontaneous increase and intensification of newborns' sucking behavior when exposed to sound stimuli that are of interest to them. This response can be measured with a pressure transducer built into a pacifier that the infants suck on. This technique is most often used to measure newborns' spontaneous preference between two types of stimuli or their ability to discriminate between them (Byers-Heinlein et al. 2010). To measure preference, infants are typically presented with two types of stimuli (e.g., two languages) in different blocks (e.g., of 1 min each). A greater number of high-amplitude sucks across blocks for one stimulus type indicates preference. Discrimination, by contrast, is often measured using a habituation/dishabituation design, whereby neonates are first habituated to one type of stimulus (e.g., one language). Once their sucking rate falls below a predefined habituation criterion, a new stimulus type (e.g., another language) is presented. If newborns increase their sucking rate after the switch, this indicates that they have noticed and responded to the change and thus can discriminate the two stimuli. Another group, the control, is tested with a switch involving a change in stimulus tokens (e.g., change in speaker) but not in stimulus type/category (e.g., no change in language). In this group, sucking rate is expected to continue decreasing after the switch, showing that the dishabituation in the experimental group is not due to low-level properties of the stimulus change.

More recently, brain imaging techniques have also been applied to investigate the newborn brain and map the neural basis of different cognitive processes. The two most commonly used techniques, both noninvasive and well-tolerated by infants, are electroencephalography (EEG) (de Haan 2013) and its less frequently used magnetic equivalent, magnetoencephalography (MEG), and near-infrared spectroscopy (NIRS) (Gervain et al. 2011, Lloyd-Fox et al. 2009). EEG measures electrical brain activity (generated by pyramidal cells) recorded from electrodes placed on the scalp. EEG has high temporal resolution (in the millisecond range), but does not provide good spatial localization due to the inverse problem, i.e., that electrical potentials travel in all directions and the signals picked up at different electrodes are a mix of the potentials originating from different sources in the brain. Mathematical techniques to solve the inverse problem and allow source localization exist but are less often used with infant data, as modeling the infant head and brain anatomy is not as advanced as adult head modeling (although see Lew et al. 2013, Roche-Labarbe et al. 2008). EEG (and MEG) data may be analyzed as (*a*) event-related potentials, that is, electrical changes in brain activity time-locked to a stimulus, or (*b*) neural oscillations, that is, neural activity in a specific frequency band (i.e., delta 1–3 Hz, theta 4–8 Hz, alpha 8–12 Hz, beta 12–30 Hz, gamma >30 Hz).

NIRS, the brain imaging modality used in this thesis project, uses red and near-infrared light of low intensity to illuminate the surface (~0.5–1.5 cm) of the cortex and thereby measure the attenuation of light caused by concentration changes in oxygenated and deoxygenated hemoglobin, which are hemodynamic indicators of local neural activity. As a result, NIRS does not measure neural activity directly, only its hemodynamic correlates. The hemodynamic response is slow, and as a result, NIRS has low temporal resolution (in the second range) but offers precise spatial localization, as it is not subject to the inverse problem [even if the resolution is low compared with magnetic resonance imaging (MRI)]. As NIRS is the technique used in all the four experiments presented in the present thesis, this technique will be described in greater details below. NIRS and EEG can also be combined (Cabrera & Gervain 2020, Telkemeyer et al. 2009, Wallois et al. 2012), as the two signals do not interfere with each other and the two types of sensors can be placed in the same headgear, typically a stretch cap. NIRS-EEG co-recording has the advantage of offering both high spatial and temporal resolution.

1.3 Newborns' universal speech perception abilities

Newborn infants have remarkable speech perception abilities. With an auditory system that is functional, but not yet fully mature (Eggermont & Moore 2012), they can flexibly start acquiring any language(s) they are exposed to. Indeed, humans are born "linguistic citizens of the world". This ability is rooted in broad-based perceptual sensitivities allowing newborns to discriminate most linguistic contrasts that appear in the world's languages.

The first task newborns face is to identify speech among the various sounds in their environment. Newborns and 2-month-old infants can indeed recognize speech and show a strong preference for it over equally complex sine wave analogs (Vouloumanos & Werker 2004). However, the category "speech" may be relatively broad at birth, roughly corresponding to primate vocalizations, as newborns show equal preference for human speech and rhesus monkey vocalizations when tested on short words versus vocalizations (Vouloumanos et al. 2010). It is only at three months that infants show a unique preference for speech over both sine wave analogs and monkey calls (Vouloumanos et al. 2010). It may be the case, though, that longer speech utterances with a full intonational contour may elicit a preference over monkey calls already at birth if prenatal experience with prosody impacts newborns' perceptual preferences.

Chapter 3 of this thesis will investigate “speech” category by measuring the newborn brain’s activations in response to speech, as well as to human sounds such as newborn cries, baby and adult laughter. If the category “speech” corresponds, at birth, to primate vocalizations, broadly defined, then brain activation to these human vocal productions will be similar to that triggered by speech. If the category is narrow and includes only speech, or even only the language heard in utero, brain activation in response to speech may already be different than responses to these human, communicative but non-linguistic stimuli.

Not only do newborns identify language in their environment, but they can also discriminate languages from one another, even if those languages are unfamiliar to them, on the basis of their different rhythms (Mehler et al. 1988, Nazzi et al. 1998, Ramus et al. 2000). Language rhythm is quantified along different acoustic dimensions, such as $\%V$, the relative proportion of vowels in the speech signal, as well as ΔC and ΔV, the variability in the length of consonant and vowel clusters, respectively (Ramus et al. 1999). On the basis of these metrics, languages can be grouped into three classes: mora-timed languages like Japanese, syllable-timed languages like French and Italian, and stress-timed languages like English and Polish. For instance, mora-timed and syllable- timed languages have relatively high $\%V$ and simple syllable structure with little variation in ΔC, while stress-timed languages have lower $\%V$ with more complex syllables, reflected in higher ΔC. [Other metrics also exist to quantify rhythm (see Dellwo 2006, Grabe & Low 2002, Loukina et al. 2011, Wiget et al. 2010).]

Importantly, rhythmic discrimination does not require familiarity with the languages. Newborns prenatally exposed to French are able to discriminate between English and Japanese. Tamarin monkeys can, too (Ramus et al. 2000), suggesting that rhythmic discrimination might be a general property of the primate or mammalian auditory system. One important implication of newborns’ ability to discriminate languages on the basis of rhythm is that infants born into a multilingual environment can immediately detect that they are being exposed to different languages, at least if those languages are rhythmically different. Bilingual newborns have indeed been shown to be able to discriminate their two languages from a third, rhythmically different language (Byers-Heinlein et al. 2010).

Newborns are thus equipped to identify speech in their environment, even in different languages. What linguistic units do they use to represent speech? They readily detect the acoustic cues correlated with word boundaries (Christophe et al. 1994). They have also been found to be sensitive to the prosodic makeup of words at the syllable level (Sansavini et al. 1997), readily discriminating words with different lexical stress patterns, that is, stress-initial (trochaic) versus stress-final (iambic). Interestingly, however, they cannot tell apart words with different numbers of phonemes if the number of syllables is the same in the two words (Bijeljac-Babic et al. 1993). They are also sensitive to the universal phonotactic constraints governing syllable structure and the distribution of consonants within syllables, known as the sonority hierarchy (Gómez et al. 2014). These results have been interpreted to suggest that the syllable is a privileged unit of representation for speech in young infants (Mehler et al. 1996). Recent EEG results (Cabrera & Gervain 2020, Ortiz Barajas et al. 2021, Ortiz Barajas et al. 2023) suggest that newborns, like adults, are indeed able to track the speech envelope (i.e., the amplitude modulation) of the speech signal, which roughly corresponds to the syllables/syllabic rate.

Young infants can also discriminate many of the phonemes appearing in the world's languages, as has been shown both behaviorally (Eimas et al. 1971 and subsequent work) and electrophysiologically (Dehaene-Lambertz & Baillet 1998). This universal discrimination repertoire is one of the hallmarks of young infants' broad-based abilities, allowing infants to learn any language they are exposed to. As with the case of rhythm, animals can also discriminate phonemes categorically at similar acoustic boundaries as humans do (Kuhl 1981, 1986), suggesting that phoneme perception is a basic ability rooted in general mammalian auditory mechanisms.

It needs to be noted, however, that most evidence is from infants 2-3 months of age and older. Newborns' phoneme representations have rarely been tested (Cheour et al. 2002, Dehaene-Lambertz & Pena 2001). One existing study suggests mature discrimination abilities, at least for native stop contrasts that are robust enough to be maintained, like in adults, even for impoverished speech signals vocoded to preserve only the lowest (<8 Hz) amplitude modulation cue (Cabrera & Gervain 2020).

Some studies, especially from non-Indo-European languages, suggest that some particularly challenging contrasts may emerge later (e.g., through refinement of existing categories) and may not be part of the initial repertoire (Mazuka et al. 2014, Narayan et al. 2010). The open questions about the initial repertoire notwithstanding, phoneme perception undergoes important perceptual reorganization on the basis of language experience during the first year of life, with infants maintaining or even gradually improving their ability to discriminate contrasts found in their native language, but in many cases losing the ability to discriminate nonnative ones (e.g., Werker & Tees 1984); for summaries see Gervain & Mehler 2010, Kuhl 2004, Swingley 2021, Werker 2018, Werker & Curtin 2005).

Newborns not only are sensitive to the sound patterns of language, but also show abilities that allow them to begin learning about language structure. They can discriminate the two universal lexical categories of functors, that is, words that mark morphosyntactic structure (e.g., *the, she, in, up*) and content words that carry lexical meaning (e.g., *door, rainbow, run, beautiful*), on the basis of the different phonological properties of the two categories (Shi et al. 1999). Newborns are also sensitive to word order and can detect the violation of the serial order in sequences of words (Benavides-Varela & Gervain 2017). They can also detect more abstract patterns, such as repetition-based regularities like ABB (e.g., “mu ba ba,” “pe na na”) or AAB (e.g., “ba ba mu,” “na na pe”), and discriminate them from otherwise similar random sequences such as ABC (e.g., “mu ba ge,” “pe na ku,” etc.) or from one another (e.g., ABB versus AAB) (Gervain et al. 2008, 2012).

How are these abilities encoded in the newborn brain? More specifically, is the brain specialized for language from the beginning? Studies conducted using NIRS and functional magnetic resonance imaging (fMRI) suggest that three-month-olds (Dehaene-Lambertz et al. 2002), full- term neonates (Peña et al. 2003), and even premature newborns (Mahmoudzadeh et al. 2013) activate approximately the same brain network as adults (i.e., the superior and middle temporal gyri, the inferior parietal cortex, and the inferior frontal gyrus, including Broca’s area) in response to language (Friederici 2012) but not to nonlinguistic controls such as backward speech. As we discuss in the next section, this specialization, in particular its lateralization (Telkemeyer et al. 2009), may already be shaped by prenatal experience. By comparing brain mechanisms at play when newborns are exposed to repetition-based regularities conveyed by syllables (linguistic condition) or by piano tones (non-linguistic condition), we will investigate if, already at birth, specific brain mechanisms are recruited for processing speech or if non-linguistic stimuli also trigger similar process (Chapter 4).

1.4 The influence of prenatal experience on speech perception abilities at birth

An increasing body of research is exploring the prenatal period, showing that fetuses already learn from the speech signal available to them during pregnancy. Moreover, several studies have shown that newborns' speech perception abilities seem to be shaped by this prenatal experience (Gervain 2015).

Starting from the 24th-28th week of gestation, the auditory system is functional (Eggermont & Moore 2012). Fetuses thus have access to different types of sounds present in the uterus: the mother's heartbeat and respiration, digestion, movements (Busnel & Granier-Deferre 1983), and voice (Petitjean 1989). It is thus during the third trimester of gestation that the fetuses first experience spoken language. Within this period, fetuses already show responses to speech. DeCasper et al. (1994), for instance, by measuring changes in fetal heart rate, showed that fetuses react differently to a children's rhyme that was often repeated in their environment compared with a novel rhyme. A recent fMRI study investigating brain activation in fetuses during the last trimester of pregnancy has found activation in the left auditory cortex in response to auditory stimuli at 33 weeks of gestation (Jardri et al. 2008). Nevertheless, the prenatal speech signal differs from the full speech signal available after birth. The exact characteristics of the speech signal heard by the fetuses in utero are not yet completely known. However, some studies, with both simulations and human and animal models using intrauterine recordings, have contributed important insights into this question (Gerhardt et al. 1992, Griffiths et al. 1994, Querleu et al. 1988). These studies have shown that external acoustic stimuli are mostly available to the fetus in the low frequencies, as the power of the higher frequencies is significantly attenuated by the tissues of the womb. More precisely, maternal tissues act as a low-pass filter around 400–600 Hz (Lecanuet & Granier-Deferre 1993). This natural filter mainly preserves the prosody of speech (i.e., the rhythm and melody of language) but drastically reduces the finer acoustic information of speech sounds, with the possible exception of some information about certain stressed vowels (Moon et al. 2013). Hence, newborn infants already perceived speech during the last three months of pregnancy, but due to the filtering effect of the womb, fetuses mainly experience speech through its prosody, and it is only after birth that the fine-grained acoustic information of the speech signal becomes available to them. An increasing body of research shows that this prenatal experience already shapes the speech perception abilities of newborns at birth.

Newborns show preferences for sounds and speech stimuli they heard in utero, indicating that they can learn from prenatal experience. One such preference newborns show is for their mother's voice compared with an unknown female voice (DeCasper & Fifer 1980, Moon 2017). Interestingly, newborns show equal preference for their mother's voice whether unfiltered (0–10 kHz) or filtered (low-pass filtered 0–1 kHz, with frequencies between 0 and 1 kHz attenuated by 6 dB and frequencies beyond 1 kHz attenuated by 48 dB), the former corresponding to the voice they hear after birth and the latter mimicking what they heard in utero (Spence & DeCasper 1987). Control newborns who were exposed to the unfiltered and filtered voice of an unfamiliar woman showed a clear preference for the unfiltered stimulus. The fact that newborns accept only the filtered voice of the mother suggests that this preference may be conditioned by infants' low-pass filtered prenatal experience. Similarly, newborns show a preference for a melody they heard repeatedly during the last trimester of gestation over an unfamiliar one (Granier-Deferre et al. 2011, Panneton & DeCasper 1986) as well as for a story that the mother had read aloud during pregnancy (DeCasper & Spence 1986).

Newborn infants can also discriminate their native language from a rhythmically different unfamiliar language and prefer listening to it over unfamiliar languages. Mehler et al. (1988) showed that French four-day-old newborns are able to discriminate utterances from their native language from utterances in another language (French versus Russian) and show a preference for the French utterances. Interestingly, when using low-pass filtered versions of the utterances, preserving prosodic information, newborns showed the same pattern of response. Newborns' preference for their native language was established by Moon and colleagues (1993), who presented English and Spanish utterances to Spanish and English two-day-old newborns. Using a preference procedure that presented auditory stimuli contingently on newborns' sucking behavior, the authors showed that Spanish newborns prefer listening to Spanish utterances and English newborns to English utterances. The case of bilingual newborns is also of interest here. Byers-Heinlein and colleagues (2010) documented that bilingual newborns whose mothers spoke both English and Tagalog during pregnancy were able to discriminate the two languages while showing equal preference for both of them. This means that, already at birth, bilingual newborns have a sufficiently detailed representation of the two languages they heard in utero to allow them to distinguish one from the other as long as they are rhythmically different.

In line with these findings, an increasing and exciting body of research suggests that this prenatal experience could shape newborns' speech perception abilities in even more specific ways and could lay the foundation for subsequent language acquisition. Specifically, several studies provide evidence that newborns learn about the distinctive prosodic features of the language(s) they heard prenatally. In four NIRS studies, Abboub and colleagues (2016) showed that, at birth, one-to- four-day-old newborns are able to discriminate well-formed prosodic sequences from ill-formed ones only when the varying feature (duration, pitch or intensity) is contrastive in the language(s) heard prenatally. Hence, French newborns were able to discriminate sequences varying in duration (short-long, typical of French prosody, versus long-short, absent from French and from most languages) (Nespor & Vogel 1986, Nespor et al. 2008) but not sequences varying in intensity (loud-soft versus soft-loud) or pitch (high-low versus low-high), two acoustic features that are present in the prosody of French but do not carry contrastive prominence (Nespor et al. 2008). Interestingly, bilingual French newborns whose other language relied on pitch to mark prosodic prominence also discriminated sequences varying in pitch. A similar sensitivity to prenatally heard prosodic patterns has also been observed in a study where newborns were able to discriminate sentences according to their emotional valence (sadness, happiness, anger, or neutrality) only when the sentences were pronounced in their native language, suggesting a prenatal shaping of the sensitivity to the prosodic variation associated with emotional expression in the language(s) heard in utero (Mastropieri & Turkewitz 1999). As prosody is mainly carried by (stressed) vowels in the speech signal, it has been hypothesized that while individual consonants are filtered out prenatally, some information about vowels may be preserved. Indeed, Moon et al. (2013) demonstrated that American and Swedish newborns who had between 7 and 75 hours of postnatal exposure to language already showed opposite preferences between the American vowel /i:/ and the Swedish vowel /y/. Both groups produced more high-amplitude sucks for the nonnative vowel. The amount of postnatal exposure did not correlate with preference, suggesting that the results reflect the influence of the prenatal period. This preference implies that the speech signal available before birth carries some vowel information, which shapes newborns’ phoneme perception abilities at birth.

Even more excitingly, the impact of prenatal learning may be so strong as to even influence production. One study found that the communicative (i.e., non-pain related) cries emitted by newborns follow the dominant prosodic contours of their respective native languages (Mampe et al. 2009). The authors found that the melodic contour of the cry patterns of French and German newborns differed and that each group showed patterns reflecting the speech prosody of their native language, that is, prominence-initial in German and prominence-final in French. While the statistical analyses in this study were criticized (Gustafson et al. 2017), in subsequent work a machine learning algorithm successfully classified newborn cries from three different languages (Manfredi et al. 2019), suggesting that the cries are sufficiently discriminable. Further work is needed to firmly establish the impact of prenatal experience on production, but if it is confirmed, it constitutes particularly strong evidence that attunement to the native language starts in utero.

In parallel, neuroimaging studies also provided evidence for prenatal modifications of the neural circuitry for speech and language processing. One issue that has received particular attention is the early hemispheric specialization of the brain for language. At three months of age, the brain activity of infants in response to sentences in their native language recruits similar left hemispheric regions as in adults, including the inferior frontal areas, superior temporal gyrus, and the angular gyrus (Dehaene-Lambertz et al. 2002). In a NIRS study conducted with Italian neonates age 2-5 days, Peña and colleagues (2003) confirmed that the left lateralization was already present at birth. They recorded the brain activity in the left and right temporal areas in response to three conditions: (*a*) Italian infant-directed utterances; (*b*) the same utterances played backward, thus suppressing the linguistic nature of the stimuli while preserving their acoustic properties; and (*c*) silence. The results showed significantly greater activation in the left temporal areas in response to speech compared with backward speech and silence, suggesting an early dominance of left regions for the processing of speech. A left hemispheric dominance was also observed for speech when compared with music (Kotilahti et al. 2009).

Interestingly, subsequent studies showed that this early cortical network, and in particular the left lateralization, was at least partly a result of prenatal experience. In their study, Peña and colleagues presented speech to newborns in their native language. In a later NIRS study, Sato and colleagues (2012) tested Japanese newborns using a similar paradigm. The newborns were presented with Japanese played forward and backward, as well as with an unfamiliar language, English, also played forward and backward. For the native language, here Japanese, the newborns showed a left hemispheric dominance for the forward as compared with the backward sentences, as in Peña et al. (2003). Surprisingly, however, no asymmetry was found between the forward and backward presentations of the unfamiliar language, English, for which responses were bilateral in both presentation directions. Another study by May and colleagues (2011) provided somewhat different results. In this experiment, English newborns listened to English and Tagalog sentences presented forward and backward. While the familiar and the unfamiliar language induced different brain responses with an overall advantage for the native language, the authors did not find a left lateralized advantage for forward-going speech in English. Rather, activation was similar in the bilateral temporal regions for both languages. The authors suggested that the absence of lateralization may be related to differences between the stimuli, which were low-pass filtered in May et al. (2011) but not in Peña et al. (2003). Subsequently, May et al. (2018) also tested nonfiltered stimuli in two NIRS studies that probed the role of experience. First, they compared forward and backward English, the native language, and Spanish, a rhythmically different unfamiliar language, and observed the expected left-lateralized advantage for forward over backward speech in the native language but comparable bilateral responses to the two directions in the unfamiliar language, similarly to Sato et al. (2012). In a second experiment exploring the specificity of the newborn brain's preparedness for language, the authors compared Spanish, unfamiliar to the infants, with Silbo Gomero, a whistled language based on Spanish. This communication system, used by shepherds on the island of Gomera, is processed in the language areas of the brain in those adults who are familiar with it, but not in those who are not. May et al. (2018) explored whether the newborn brain starts out with the potential to process whistle as language and loses this ability in the absence of experience, or whether the newborn brain is more specifically tuned to speech and learns to process whistle as language only when exposed to it. The results pointed to the second option. Newborns showed an increased and left-lateralized response to forward Spanish but essentially no response to Silbo Gomero, suggesting that the brain's initial specialization does not extend to whistle. Interestingly, the unfamiliar language Spanish gave rise to a bilateral response with no forward advantage when contrasted with the native language English, but a left-lateralized response with an advantage for forward speech when Spanish was compared with Silbo. This implies that context effects may also play a role in how the newborn brain processes speech. An additional factor that influences the localization of the responses may be the acoustic properties of the stimuli, in particular for the nonnative language. May et al.'s (2011) results with low-pass filtered stimuli also point in this direction. Additionally, while observing the same left-lateralized forward advantage for the native language, French, as in other studies, Vannasing et al. (2016) found right-lateralized activation for the forward unfamiliar language, Arabic. The lateralization issues notwithstanding, all studies found differences in the brain responses to the native language as compared with unfamiliar languages, supporting the view that prenatal experience shapes the early language cortical network.

One interesting question is: What properties of the prenatal signal drive the brain specialization for the native language? Specifically, is it simply having experience that matters, or does the low-pass filtered nature of the prenatal signal play a role? While we do not currently have a definitive empirical answer to this question, one NIRS study (Bartha-Doering et al. 2019) comparing responses to forward and backward speech in the native language in full-term newborns and in preterm newborns tested at term age suggests that the quality of the experience may be crucial. In this study, full-term newborns showed a left-lateralized advantage for forward speech, like in all previous studies with nonfiltered stimuli, but preterm infants showed no such advantage. Since they were otherwise healthy and they were tested at term age (i.e., they had as much experience with language as their full-term peers at test), but this experience was full-band rather than low-pass filtered, this result suggests that missing out on the prenatal signal may be detrimental. Other studies with older infants who were born preterm seem to support this view, as these infants typically show language delays for aspects of language that are related to prosody (e.g., Peña et al. 2010), but less often show delays for linguistic patterns that are only experienced postnatally, such as phonotactics (e.g., Gonzalez-Gomez & Nazzi 2012). Phoneme perception itself may also be delayed (Peña et al. 2012). Since some stressed vowels may be perceived in utero, further research is needed to better understand how the perception of different phoneme classes patterns (i.e., delayed or not) and assess whether the observed pattern of results is compatible with the current hypothesis. This hypothesis needs further empirical confirmation, but existing data seem to suggest that as far as prenatal experience is concerned, less is more, that is, the impoverished, prosody-only nature of this signal may be just the right kind of input for the developing auditory and speech perception abilities of the fetus.

Newborns' ability to recognize their native language at birth can be seen as fundamental to focus on this specific auditory signal in order to acquire language. Chapter 5 will investigate how this crucial capacity may be impacted by reduced experience with speech. To do this, we conducted a study to measure the brain activation of infants with hearing loss in response to their native language as well as to an unfamiliar, rhythmically different language. What is the minimum input necessary for newborns to recognize their native language at birth?

1.5 The current thesis

Newborns already show impressive speech perception abilities that allow them to acquire their native language(s). At birth, they already show evidence that they have learned from prenatal experience, being more sensitive to their native language(s) compared to unfamiliar, rhythmically different language(s). The native language triggers a left-lateralized and/or increased brain activation, similar to adults, demonstrating the neonatal brain's preparedness for language. But what is "speech" for newborns: a narrow category including only (their native) language(s) or a broader category that narrows down to human language only with experience during the first months of life?

1.5.1 What is speech for newborns?

Existing studies suggest that infants and newborns show a bias for listening to speech compared to other auditory stimuli (Colombo & Bundi, 1981; Glenn, Cunningham and Joyce, 1981; Vouloumanos & Werker, 2004; 2007; Vouloumanos et al., 2010). Infants prefer listening to speech over white noise (Colombo and Bundi, 1981) and to a woman singing a song a cappella compared to listening to this same song played by musical instruments (Glenn, Cunningham and Joyce, 1981). To compare speech with an acoustically better matched auditory stimulus, Vouloumanos and Werker (2004) investigated infants' preference for listening to speech over complex non-speech sine wave analogs. The infants tested, who varied in age (2.5, 4.5 and 6.5 months of age), all showed longer looking times in the speech trials than in the sine wave analogs. This pattern of results was replicated subsequently with neonates (Vouloumanos and Werker, 2007). These various findings suggest that infants, and even newborns, present a listening bias for speech compared to other auditory stimuli, even auditory stimuli that have the same acoustic complexity as speech.

However, when newborns were presented with speech and rhesus monkey vocalizations, the listening bias for speech was no longer found (Vouloumanos et al., 2010). Newborns preferred to listen to monkey calls over complex sine wave speech analogs and showed no preference between speech and monkey calls. Thirty neonates were presented with speech (non-sense words) and monkey vocalizations. They sucked similarly strongly during the presentation of both stimuli, indicating no preference for one stimulus over the other, but they preferred listening to both speech and monkey calls over complex synthetic sound sine wave analogs. This pattern changed with speech experience, as three-months-olds did show a preference for speech over monkey vocalizations. This set of results suggests that, at birth, listening preferences may be quite broad, comprising both human speech and other primate vocalizations, while after at least three months of experience with speech, the preferences narrow down to speech. The same pattern has been found in the visual domain looking at newborns' preferences for human faces over monkey faces: results show that newborns similarly process human and monkey faces and do not show any preference for one stimulus over the other, despite being able to discriminate them (Di Giorgio et al., 2012).

Is this also the case for speech? Speech, in its acoustic, phonemic, lexical, morphological and syntactic complexity is unique to humans. The question thus remains what constitutes the category "speech" at birth. Is it a broad category comprising all primate vocalizations? This is what Vouloumanos and colleagues' results (2010) suggest. However, the speech stimuli used in those experiments were non-sense words, i.e. very simple, short, not ecological speech stimuli. In particular, they lacked the rich prosodic modulation typical of full sentences, which constitute newborns' prenatal experience. The reason for this was that monkey calls are inherently short and acoustically less complex, so the authors aimed to match the speech stimuli in duration and complexity. This, by necessity, entails that the speech stimuli used were not particularly ecological and lacked the richness of connected speech. We hypothesize that if the speech condition had been composed of full sentences, the authors may have found a preference for speech over monkey vocalizations.

To better understand what constitutes the category “speech” at birth, the current thesis thus tests newborns’ brain responses to speech and various other auditory stimuli that are speech-like in various ways. We decided not to compare speech with monkey calls given the inherent differences in duration and complexity between the two types of stimuli, as discussed with respect to the Vouloumanos et al. (2010) study. Rather, we decided to use vocalizations that are produced by humans, and which, albeit non-linguistic, have communicative functions.

Specifically, we chose two stimuli, cries and laughter, that are communicative, produced by the human vocal tract and that even young infants are able to produce. In particular, newborns’ cries are very relevant as they are sounds that the newborns themselves can produce and that are perceived as communicative by the caregivers. Moreover, recent studies suggest that newborn cries may show a developmental continuum with language, as they already convey the prosody of the native language heard in utero (Mampe et al., 2009; Manfredi et al., 2019). In Experiment 1, we thus ask whether newborns’ cries elicit brain activations similar to spoken language (Chapter 3)?

As a second comparison, we used laughter. Laughter is an auditory stimulus produced by the human vocal tract, but, unlike cries, it does not carry prosody and has a positive rather than a negative emotional valence (Experiment 2, Chapter 3).

These first two experiments aim to better understand what is identified as speech at birth, i.e. if speech triggers a specialized brain processing in newborns or if different communicative vocalizations produced by humans are first perceived as similarly relevant in a broad category that later narrows down after exposure to speech.

1.5.2 Does speech trigger specific brain mechanisms compared to other auditory stimuli?

We know that, already at birth, newborns are able to learn from the linguistic input, being able, for instance, to extract simple repetition-based rules when presented with syllables (Gervain et al., 2008; 2012). Is this ability specific to speech as a privileged vehicle of language or do these learning mechanisms first apply to various auditory stimuli, whether linguistic or not? In other words, do newborns relate the category speech to the abstract properties of language or does a broader category of (auditory) stimuli trigger abstract structural computations, e.g. rule extraction? To investigate this, in Chapter 4, we compared newborns' ability to extract a repetition-based rule from speech (pseudowords) and from a non-linguistic input (sequences of piano tones).

1.5.3 How is the category speech perceived and represented in atypical development, i.e. infants with hearing loss?

At birth, newborns are already sensitive to the prosodic features of the language(s) they heard in utero (Abboub et al., 2016). They are able to discriminate it from an unfamiliar rhythmically different language (Mehler et al., 1988; Nazzi et al., 1998; Ramus et al., 2000) and the native language triggers left-lateralized brain activation (Peña et al., 2003, Sato et al. 2011), suggesting prenatal learning. When prenatal experience with speech is disrupted, the brain processing of the native language can be impacted. For instance, Bartha-Doering and colleagues (2019) did not find left lateralization in response to the native language in preterm babies, compared to full-terms. In newborns and infants with hearing loss, both the pre- and the post-natal experience with speech is disrupted. How does this affect basic speech perception abilities, i.e. the recognition and discrimination of the native language from other languages based on prosodic processing? To investigate this, we tested the rhythmic discrimination of languages in infants with hearing loss, using a paradigm commonly used in the literature (May et al., 2011; 2017; Sato et al., 2011), comparing infants' brain responses to the native language and to an unfamiliar, rhythmically different language (Experiment 4).

2 Functional Near-Infrared Spectroscopy (fNIRS): measuring how the newborn brain processes speech

2.1 The basic physical principles of NIRS

In this thesis, the four experiments presented investigate newborns and young infants' brain responses to various auditory stimuli, all using functional near-infrared spectroscopy (fNIRS). This brain imaging technique, also called optical topography (OT), uses infrared and near-infrared light to detect changes in blood oxygenation in focal brain areas related to activation. Neural activity is accompanied by changes in blood oxygenation; when a brain area is activated by a stimulus, this leads to an increase of blood flow, which brings a transient excess in oxygen. What is measured by NIRS (similarly to functional Magnetic Resonance Imaging, fMRI) is this hemodynamic response, which appears a few seconds after the onset of stimulation (Figure 2.1). This is thus an indirect measure of neuronal activity. Most commercially available NIRS machines use continuous wave (CW) monochromatic red and/or near-infrared light at two wavelength and measure

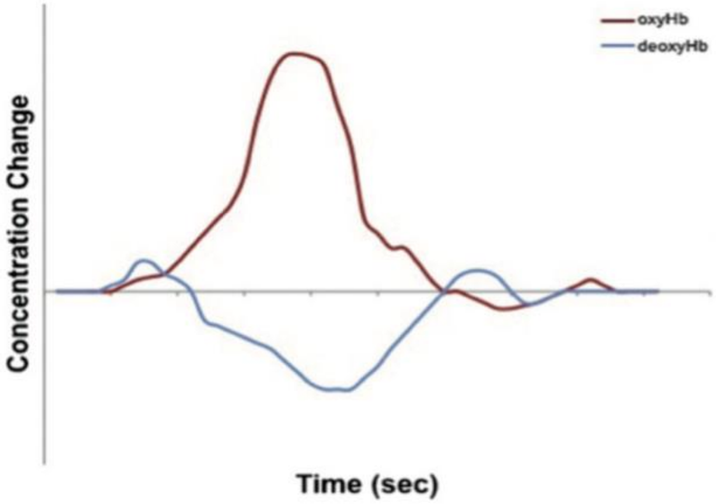


Figure 2.1 : A typical hemodynamic response in adults. The red line represents the concentration in oxygenated hemoglobin (oxyHb) and the blue line represents the concentration in deoxygenated hemoglobin (deoxyHb). The stimulation is given at t=0. Typically, there is first a small increase in deoxygenated hemoglobin followed by a large increase of the oxygenated hemoglobin while the concentration in deoxygenated hemoglobin is decreasing. This hemodynamic response appears a few minutes after the stimulus and then goes back to baseline. In the newborns populations, this hemodynamic response is slower.

the absorption of this light by the cortex. When the light goes through the tissues, one part of it is absorbed, one part is scattered and one last part is transmitted, i.e. it goes through the tissue without impacting it. The Beer-Lambert law allows us to determine the concentration of oxygenated and deoxygenated hemoglobin from the absorption of light (A), where (ε_λ) is the absorption coefficient and (c) the concentration of the medium, (λ) is the continuous wave used, (l) is the distance covered, (I) the intensity of the transmitted light after passing through the medium and (Io) the intensity of the incidental light:

$$A = -\log\left(\frac{I}{I_0}\right) = c \times \varepsilon_\lambda \times l \quad (1)$$

Light passes through several tissues before reaching the cortex (skin, skull, cerebrospinal fluid, etc.). To take into account the complexity of these different tissues and the dispersion of the light caused by these different media, the Beer-Lambert Law needs to be modified in order to consider the dispersion or scatter (G) and the non-linear trajectory of the light within these biological media expressed by the differential path length factor (DPF):

$$-\log\left(\frac{I}{I_0}\right) = (c \times \varepsilon_\lambda \times l \times DPF) + G \quad (2)$$

Since the scatter and the differential path length factor are not known, but are constant and can thus be dropped when change between two time points is considered, we can use the modified Beer-Lambert Law to calculate the relative values of the concentration of oxygenated and deoxygenated hemoglobin, and thus to measure the changes in blood oxygenation (Bouchon et al., 2013).

2.2 NIRS hardware

NIRS uses a coupling of an emitter (a source) and a detector of light. This pair forms a channel. The distance between the source and the detector determines how deep the light penetrates into the brain (Figure 2.2). The greater the distance, the deeper the light goes into the cortex. Since this penetration is not deeper than 3cm from the scalp surface with the usual 2-4cm separation used in most commercially available machines, which is about 0.5cm into the cortex in adults and about 1-1.5cm into the cortex in infants, NIRS can only measure from the superficial layers of the cortex. This allows the investigation of the brain areas crucial for language and auditory processing for instance, but does not include the deepest brain layers. The optimum distance between the source and the detector depends on the participants' age. Newborns have thin skin and tissues, thus a relatively short distance allows to reach the entire thickness of the cortex, while the same distance in an adult participant would only allow to reach the superficial tissues. The difference is about three times greater for newborns than for adults; e.g. a distance of three centimeters between the source and the detector allows a penetration of the light of 10 to 15 millimeters in the cortex of newborns, while only of 3 to 5 millimeters in adults.

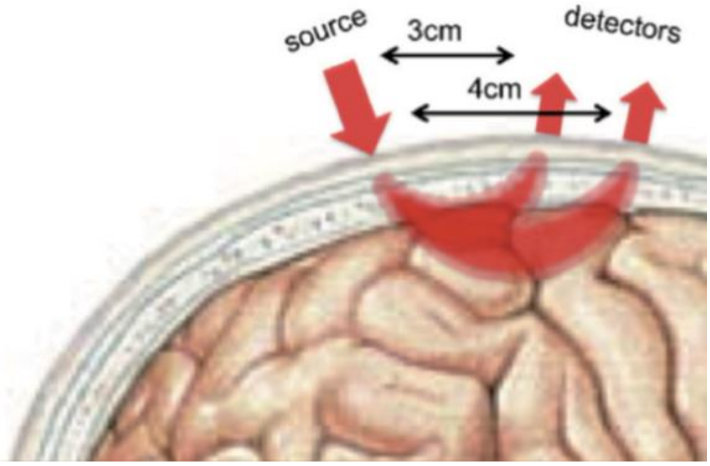


Figure 2.2 : Comparison of the light penetration into the cortex depending on the source-detector distance.

2.3 Use of NIRS with developmental populations

Near-infrared spectroscopy has been used with developmental populations for almost thirty years now (Meek et al., 1998; Aslin and Mehler, 2005), as it is well-suited for psychological research (Ferrari et al., 2014) and especially for the investigation of language and speech processing in young populations (Obrig et al., 2010; Gervain et al., 2011). NIRS is well-suited for newborns and young infants for various reasons: it is a non-invasive technique, posing no safety concerns, it is portable making it very easy to use in different environments (hospital, maternities, participants' home). NIRS is also less sensitive to motion artefact than other techniques like fMRI; thus there is no need to immobilize the participants, which is a great advantage when testing developmental populations. Finally, this technique is totally silent, allowing researchers to use it while recording brain activation in response to auditory stimulation, which is fundamental to investigate brain responses to spoken language. Indeed, NIRS studies have brought significant findings to the language acquisition field, highlighting the presence of a left-lateralized pattern of activation in response to the native language already at birth (Peña et al., 2003) or newborns' ability to extract regularities from the speech input (Gervain et al., 2008; 2012).

The NIRS community in developmental research keeps growing, improving the technique and moving ever closer to the practice of a rigorous, open science (Kelsey et al., 2023).

3 Newborns' perception of speech and other human sounds

3.1 *Experiment 1:* Are newborns' cries already language?

3.1.1 Introduction

Several studies have suggested that newborns and young infants prefer speech over other auditory stimuli (Colombo & Bundi, 1981; Glenn, Cunningham and Joyce, 1981; Vouloumanos & Werker, 2004; 2007; Vouloumanos et al., 2010). Four-month-old infants prefer to listen to continuous speech produced by a woman, compared to continuous unfiltered white noise (Colombo and Bundi, 1981) and 9-month-olds prefer to listen to a woman singing a song a cappella compared to listening to this same song played by three different solo musical instruments (Glenn, Cunningham and Joyce, 1981). Considering that the control conditions in these studies were quite far from the acoustics of speech, Vouloumanos and Werker (2004) investigated infants' preference for speech over an auditory stimulus that was more similar to speech both in the spectral and the temporal domains. For this, they created complex non-speech analogs composed of time-varying sinusoidal waves that preserved variations of energy present in natural speech (i.e. the fundamental frequency and the first three formants). Three groups of infants of different ages (2.5, 4.5 and 6.5 months of age) were exposed to speech stimuli (nonsense words) and to non-speech analogs, while their preference was measured using an infant-controlled sequential looking preference paradigm. Results showed that all the infants looked longer at the speech trials than at the non-speech ones. In a subsequent study (Vouloumanos and Werker, 2007), the authors carried out this same study, but this time with neonates, to determine if this bias for speech was already present at birth. Twenty-two neonates were exposed to the speech and non-speech stimuli, while their responses were measured using the high-amplitude sucking (HAS) procedure. The authors found that newborns sucked significantly more when they were listening to speech compared to when they were listening to the complex sine-waves analogs. This set of results thus shows that infants display a preference for speech already from birth, compared to artificial non-speech stimuli.

However, when speech was contrasted with rhesus monkey vocalizations, newborns no longer showed a preference for speech (Vouloumanos et al., 2010). Thirty neonates were presented with speech (non-sense words) and monkey vocalizations. They sucked similarly strongly during the presentations of both stimuli, indicating no preference for one stimulus over the other, but they preferred both speech and monkey calls over sine-wave analogs. Three-month-old infants, by contrast, already showed a preference for speech over monkey calls. Taken together, these results suggest that, at birth, neonates listening preferences may be quite broad, integrating both human speech and other primate vocalizations, while after three months, their preference narrows down to speech.

The same pattern has been shown in the visual domain. At birth, newborns show preference for faces over other complex visual stimuli (Fantz, 1961; Goren, Sarty & Wu, 1975; Johnson & Morton, 1991; Valenza, Simion, Macchi Cassia and Umiltà, 1996; Macchi Cassia, Turati and Simion, 2004). However, when newborns are presented with images of human faces and of monkey faces, they show no preference for one type of stimulus over the other (Di Giorgio et al., 2012), although they are able to discriminate the two categories. Moreover, as for human faces, newborns' preference is disrupted for inverted, upside down monkey faces. Thus, at birth, newborns process human and monkey faces in a similar way. This could be explained by the fact that humans are primates and our faces have roughly the same shape and same configuration as those of other primates.

Can a similar narrowing also be observed for speech or is the category "speech" already more specific at birth? It has been shown that some primates have sophisticated vocal communication (for a review about chimpanzee vocal communication, see Slocombe et al., 2022), with the possibility of combining, at least to some extent, different units in order to convey different meanings (Coye et al., 2016; Bortolato et al., 2022), but speech, in its acoustic, phonemic, lexical, morphological and syntactic complexity is unique to humans. The question thus remains what constitutes the category "speech" at birth. Is it a broad category comprising all primate vocalizations?

This is what Vouloumanos and colleagues' results (2010) suggest. However, the speech stimuli used in those experiments were non-sense words, i.e. very simple, short, non-ecological speech stimuli. In particular, they lacked the rich prosodic modulation typical of full sentences, which constitute newborns' prenatal experience. The reason for this was that monkey calls are inherently short and acoustically less complex, so the authors aimed to match the speech stimuli in duration and complexity. This, by necessity, entailed that the speech stimuli used were not particularly ecological and lacked the richness of connected speech. We may thus hypothesize that if the speech stimuli contrasted with monkey vocalizations were more ecological and natural, in particular if they were full sentences, we may find a preference for speech already at birth.

To better understand what stimuli constitute the category "speech" at birth, we decided to test newborns' brain responses to speech compared to other human-produced sounds. The first such sound we tested was crying. Newborns' cries are highly relevant for this question as they are sounds that newborns themselves are able to produce and that are perceived as communicative by the caregivers. Moreover, recent studies suggest that newborns' cries may already be related to language (Wermke and Mende, 2009), as they already convey the prosody of the native language heard in utero (Mampe et al., 2009; Manfredi et al., 2019) (Figure 3.1). Mampe and colleagues (2009) acoustically compared prenatally German-exposed and French-exposed newborns' communicative cries. They operationally defined a communicative cry as a cry characterized by a rising-and-then-falling contour in a single melodic arc obtained in a situation where no obvious stressor or other source of distress was present. The authors investigated the pitch and intensity variations of cries emitted by prenatally German- and French-exposed newborns. They calculated the time point at which the maximum pitch (maximum fundamental frequency, F0) and the maximum intensity were reached along the melodic arc. They found that French newborns preferentially produced cries with a rising melodic contour, i.e. pitch and intensity maxima were reached during the second half of the cry, while German newborns showed the opposite pattern, with a falling melodic contour (i.e. the maximum intensity and pitch were reached early in the cry). These patterns are congruent with the general melodic patterns of the two languages: French intonation is characterized by prominence at the end of prosodic units, while German intonation is generally prominence-initial. These results suggest that prenatal experience with the native language would be so important that it could influence not only perception, but also newborns' first vocal productions. However, these intriguing results have been challenged by a subsequent study (Gustafson et al., 2017) in which the authors highlighted issues with the statistical analyses used in Mampe and colleagues (2009). Specifically, as each baby has his/her cry signature, which means that all cries of one specific baby share some acoustic features, cries from one baby should be statistically analyzed as linked and cannot be treated as independent, as it was the case in Mampe and colleagues 's (2009) paper. In their study, Gustafson and colleagues analyzed the cries of neonates born in a Mandarin Chinese linguistic environment, and of neonates born in an American English environment. They analyzed the data in two ways (i) by treating each cry as independent, even if coming from the same baby or (ii) by treating the cries of the same baby as nested. With the first analysis, they found several significant differences between the two language groups. However, when they took into account the nestedness of cries in participants, no significant differences were found any more. These results put into question the reliability of Mampe and colleagues' results. In a more recent study (Manfredi et al., 2019), researchers used a voice analysis tool and an automatic classifier to test whether cries of newborns from three different linguistic backgrounds (French, Arabic and Italian) can be correctly classified in three groups, based on acoustic analysis. Seven thousand five hundred cries were given to the machine and the classifier was able to correctly classify the cries in the three different linguistic groups with 95% accuracy. These latter results suggest, similarly to Mampe and colleagues (2009), that newborns' cries are already shaped by the native language heard in utero. It needs to be noted, however that a machine learning classifier may rely on features other than prosody to classify cries, and some of these features (e.g. the quality of the recordings, or the type of background noise) may be completely irrelevant for the broader question of whether newborns cries already resemble the prosodic contours of the prenatally heard language.

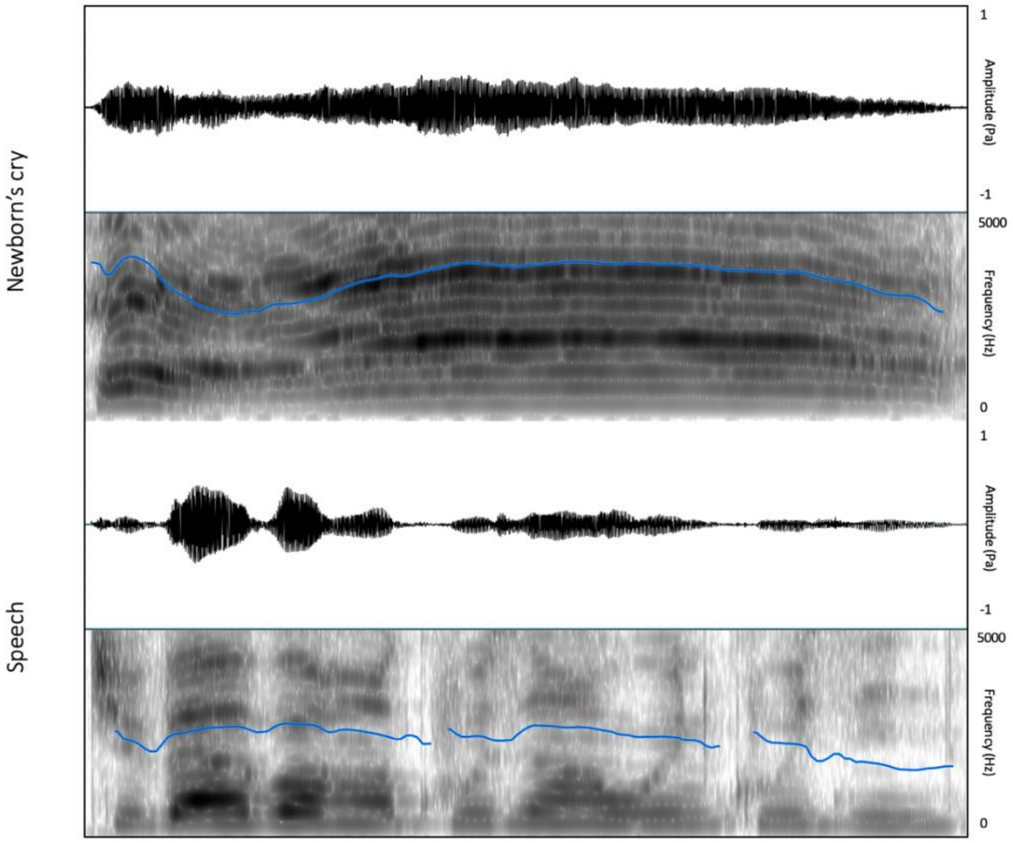


Figure 3.1 : Waveform (upper panels) and spectrograms (lower panels) of a French newborn's cry and a French sentence produced by a female native speaker. The time scale on the x axis is the same for both. The blue line represents the pitch contour.

The scarce and mixed results on this topic do not license strong conclusions and further studies are needed to determine more clearly the influence of the native language on the vocal productions of newborns.

One way to investigate the linguistic status of the newborns' cries is to compare the brain responses to cries to those elicited by speech. Does listening to cries elicit brain activation in the speech processing network of the newborn brain? To investigate this, we presented newborns with sentences in their native language, French, and with cries of prenatally French-exposed newborns, while their brain activity was measured using fNIRS in the frontal, temporal and parietal regions.

When listening to newborn cries, newborns are listening to an auditory stimulus that they themselves can produce. The motor theory of speech perception (Liberman, 1967) claims that there is a strong link between speech perception and production. For instance, Stetson (1928) wrote: *"Speech is rather a set of movements made audible than a set of sounds produced by movements."* In other words, according to this theory, when hearing speech, perception retrieves the articulatory gestures that generated the sounds we hear.

The listener uses his/her own experience as speaker to identify, through hearing, the articulatory gestures of the person speaking. This theory has limited evidence in developmental research. Some studies, however, suggest that there are sensorimotor influences on speech perception even in infants (Yeung et al., 2013; Bruderer et al., 2015). Bruderer and colleagues (2015) manipulated the placement of the articulators in English-learning 6-month-old infants while they were exposed to sound contrasts differing in tongue tip position. The authors used a Hindi phonemic contrast between the dental /d̪/, produced by placing the tongue tip at the back of the teeth, and the retroflex /ɖ/ produced by curling the tongue tip. In the first experiment, the authors replicated previous findings by showing that 6-month-old English-learning infants discriminate this non-native Hindi contrast. In a second experiment, a new group of infants was exposed to the exact same sound contrast, while they had a teething toy in their mouth that hinders the movement and control of the tongue tip. In this case, no discrimination of the contrast was found. In a last experiment, the teething toy given to infants had a shape that did not impede tongue tip movement and control. In this case, infants were able to discriminate the non-native contrast. These results suggest a sensorimotor influence on speech perception in 6-month-old infants: temporary inability to produce a sound seems to impede the auditory perception of this sound.

If there is indeed a link between perception and production already in young infants who do not yet produce a broad repertoire of language sounds, this may support the hypothesis that newborns process cries in a privileged manner, as these are communicative sounds they themselves can produce.

We thus hypothesized that sentences in the native language will activate the left (or possibly the bilateral) temporal, i.e. auditory, cortices, as observed in previous studies (e.g. Peña et al. 2003, Sato et al. 2011). If cries are also processed as if there were speech-like, then we expect to find similar activation for cries to that observed for speech.

3.1.2 Participants

Twenty-five healthy full-term newborn infants, who had been exposed to French prenatally, were included in the final analysis (15 females; mean age: 2.05 days; age range: 1–4 days, mean weight: 3320g, weight range: 2860-4100 g, Apgar score 5 minutes after birth: min. 8). Forty-five additional infants were tested but excluded from analysis due to poor data quality (11), fussiness or crying (22), less than 80% exposure to French during the pre- and postnatal period (5) and technical error (7). Recruitment was carried out as follows: the eligible newborns were first selected by an examination of the birth record. A baby was considered eligible if he/she was born at least at 37 weeks of gestation, had a minimum birth weight of 2700 grams, a minimum Apgar score of 8 five minutes after birth, had no known metabolic, neurologic or other health concerns. Once the eligible babies were selected, the experimenter talked to the parents about study participation. Parents received an information leaflet (procedure, technique used, participants' rights, data protection etc.). If the parents wished to participate, an appointment was made at a time that did not interfere with medical care. Infants were tested in their own hospital room, with at least one of the parents present at all times. All parents of the newborns tested gave written informed consent before the study, which was approved by the CPP Sud Est V ethics committee (N° ID RCB: 2018-A01136-49).

3.1.3 Stimuli

The stimuli were composed of cries recorded from 10 French newborns (none of which participated in the study) and of spoken French sentences produced by 10 adult female French native speakers. Each cry was matched in length with an utterance (mean cry length 1095 ms, SD 0.49, mean utterance length: 1094 ms, SD 0.49, W = 5029, $p = .944$), and their intensity was normalized. Only female, and not male, voices were chosen as female fundamental frequencies are generally higher than those of males, thus matching better with the high frequency of the newborns cries. The experiment consisted of a total of 20 blocks, 10 blocks of cries and 10 blocks of speech. Every block of cries comprised 10 cries, each produced by a different newborn. Similarly, every block of speech comprised 10 utterances, each produced by a different French speaker. No utterance or cry was repeated during the study. The order of presentation of the blocks was intermixed and counterbalanced, such that no more than two blocks of the same condition (cry, speech) occurred consecutively, and varied across participants. Each infant listened to the stimuli in a different order. Stimulus presentation within blocks was randomized across participants. Blocks were separated by silent intervals jittered between 25 and 30 s. Within blocks, cries/sentences were also separated by silences jittered between 500 and 1500 ms, yielding blocks of approximately 20 s (range 19.43-20.41 s). The whole experiment lasted about 16 min (Figure 3.2). [The study was conducted in collaboration with Irene de la Cruz-Pavía (Deusto University, Bilbao, Spain), who created the stimuli.]

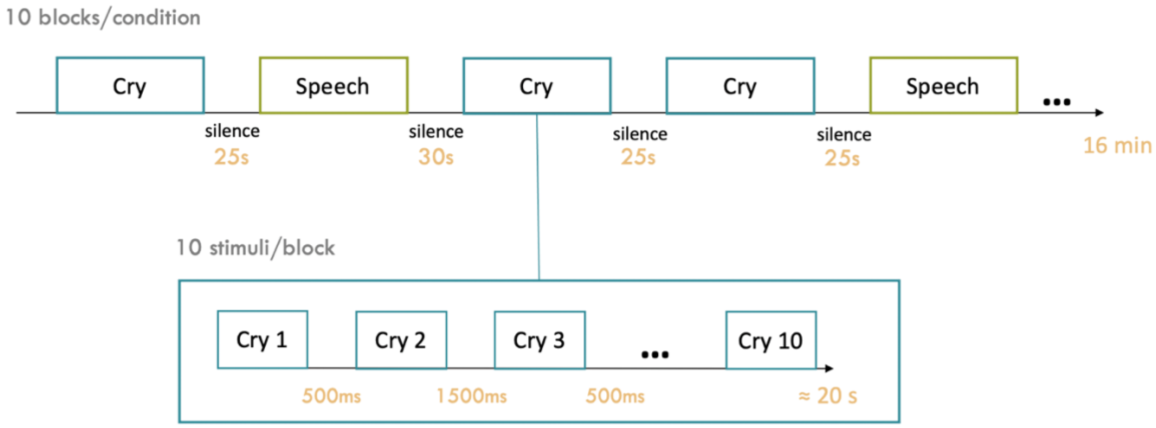


Figure 3.2 : Design of the study. The upper boxcar shows the organization and timescale of the entire experiment; the lower boxcar illustrates the organization of a block.

3.1.4 Procedure

Newborns were tested in their room at the maternity ward of the Hôpital Necker-Enfants Malades (Paris, France), in the presence of at least one of their parents. During the study, infants lay in their cribs asleep or in quiet rest, while their brain responses were measured with a NIRx NIRScout 16–16 machine (source-detector separation: 3 cm; pulsated LED lights at two wavelengths of 760 nm and 850 nm; sampling rate: 15.625 Hz). The optical probes were inserted into a stretchy cap (EasyCap, Brain-Products GmbH, Germany) placed on the infants' head using surface landmarks (nasion, and the preauricular points), covering the language areas in the bilateral temporal, frontal, and parietal cortices (12 channels/hemisphere, see Figure II-XX). The size of the cap was chosen based on each newborn's head circumference.

To localize the 24 channels, we projected their position relative to surface landmarks onto the cortical surface of 3D newborn anatomical head scans following Lloyd-Fox et al. (2014) and Shi et al. (2011), as in previous studies from our laboratory (Abboub et al., 2016; Martínez-Álvarez et al., 2022). Channels 1, 2, 4, 5 on the LH and 13, 14, 15, 16 on the RH were on average positioned over the frontal area, channels 3, 6, 8, 11 on the LH and 17, 19, 22, 24 on the RH over the temporal area, channel 9 on the LH and 21 on the RH are positioned in the frontier between the temporal and parietal areas, and channels 7, 10, 12 on the LH and 18, 20, 23 on the RH were on average positioned over the parietal area (see Figure 3.3).

Stimuli were presented using E-Prime, and delivered through two speakers placed at an angle of 30° on the two sides and approximately 1 m from the newborn's head, and elevated to the height of the crib. The stimuli were presented at a conversational intensity of around 65dB. The computer running E-Prime sent time stamps to the NIRS machine.

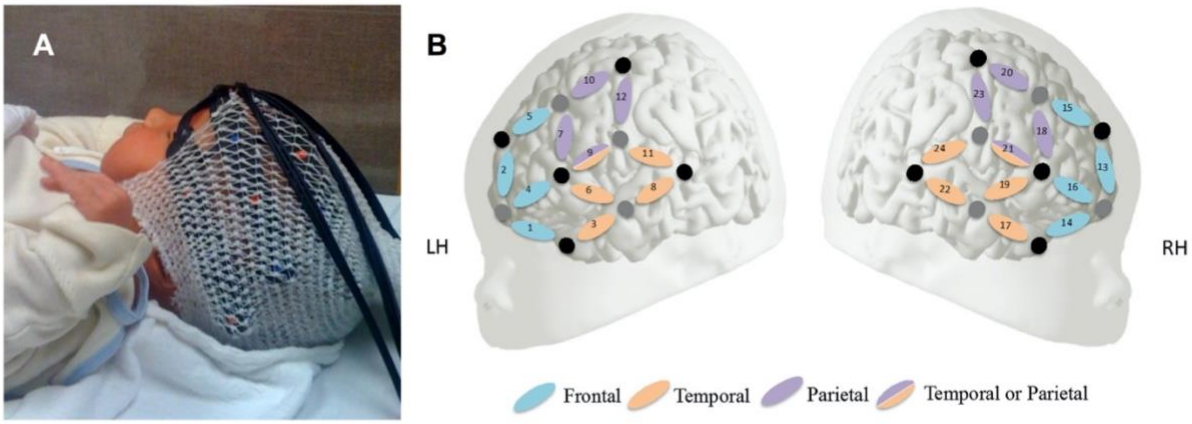


Figure 3.3 : The position of the NIRS headgear. (A) A study participant with the headgear. (B) Configuration of light sources and detectors as well as NIRS measurement channels used in the study overlaid on a schematic newborn brain. Grey circles indicate detectors, while black circles indicate sources. Localization analysis was conducted as described in Abboub et al. (2016), localization figure adapted from Abboub et al. (2016).

3.1.5 Data analysis and statistics

Changes in the concentration of oxygenated hemoglobin (oxyHb) and deoxygenated hemoglobin (deoxyHb) were calculated from the absorption of near-infrared light as metabolic indicators of neural activity. Data were band-pass filtered between 0.01 and 0.7 Hz. Movement artifacts, defined as concentration changes larger than 0.1 mmol x mm over 0.2 msec, were removed by rejecting block-channel pairs where artifacts occurred. For the non-rejected blocks, a baseline was linearly fitted between the means of the 5 sec preceding the onset of the block and the 5 sec starting 40 sec after the onset of the block (25 sec of stimulation plus 15 sec of resting period). This pipeline was used in a large number of previous studies in our lab and validated on actual and synthetic data recently (Gemignani & Gervain 2021).

Statistical analyses were carried out over the concentration changes of oxyHb and deoxyHb. We conducted a cluster-based permutation analysis (Maris & Oostenveld, 2007) to identify the existence of clusters of channels in which a significant difference in concentration changes between the two conditions were found. We also compared each condition to the baseline. This analysis allows to define regions of interest (ROIs) and time windows of interest in a non-arbitrary and data-driven manner, while preventing multiple comparisons (Mahmoudzadeh et al., 2013; Benavides-Varela & Gervain, 2017). The definition of spatial clusters was anatomically constrained, as determined by a localization analysis (Abboub et al., 2016; Figure II-2). One thousand permutations were conducted under the null hypothesis, for both oxyHb and deoxyHb.

3.1.6 Results

Figure 3.4 represents the grand average hemodynamic response in the 24 channels for oxyHb and deoxyHb concentration changes.

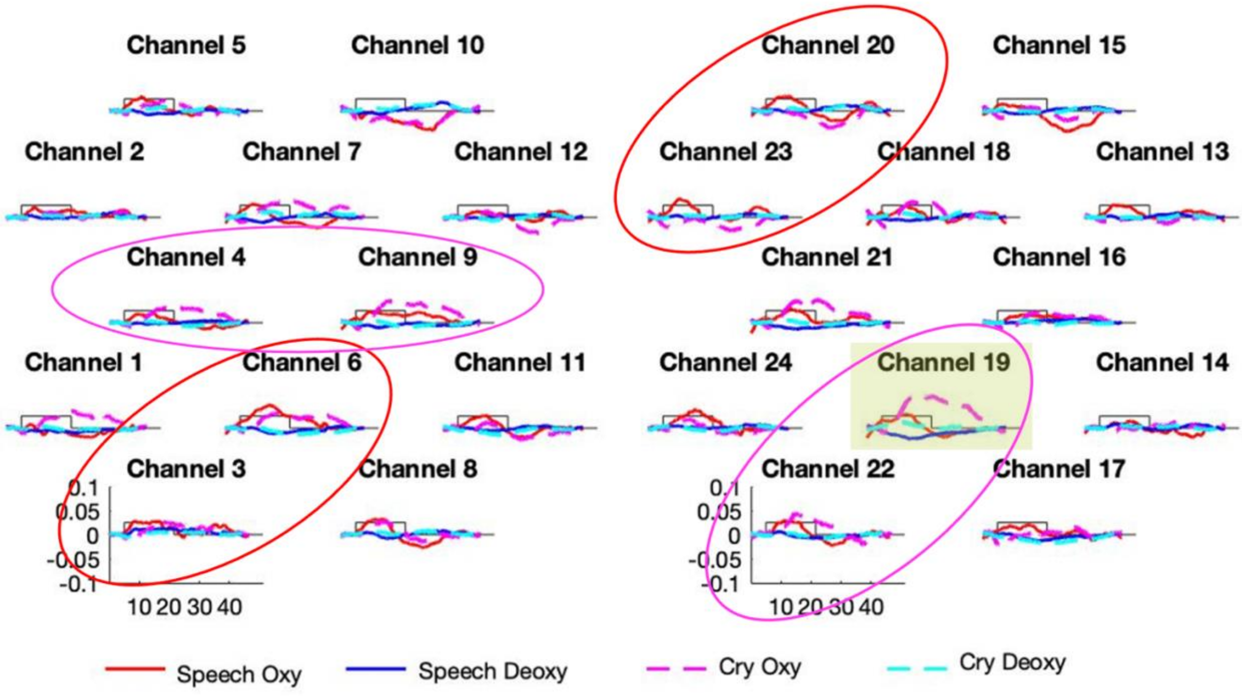


Figure 3.4 : The grand average hemodynamic response. Channels are plotted following the placement displayed in Figure 3.3. The x-axis represents time in seconds; the y-axis shows the concentration in mmol x mm. The rectangle along the x-axis indicates the time of stimulation. OxyHb and deoxyHb concentration changes in response to the Speech condition are shown in red and blue respectively (continuous lines). OxyHb and deoxyHb concentration changes in response to the Cry condition are shown in magenta and cyan, respectively (dashed lines). The results of the cluster-based permutation analysis on oxyHb are plotted as follows: the channel in which a significantly different brain response was found between the two conditions is highlighted by a green rectangle. The clusters in which the speech condition (red) or the cry condition (magenta) triggered a significantly greater brain activation than the baseline are shown by colored circles.

OxyHb.

Speech vs Cry. The cluster-based permutation analysis revealed a right temporal channel (ch19) (p=0.01) and a right parietal channel (ch23) (p<.001) where the response to the two different conditions (speech vs cry) was significantly different. Specifically, channel 19 shows a canonical and much greater activation for cry than for speech, while in channel 23, the response to speech is greater than to cry.

Speech vs Baseline. The analysis revealed a left temporal cluster of two channels (3 and 6) (p<0.001) and a right parietal cluster of two channels (20 and 23) (p<0.001) where the response to the speech stimuli was significantly greater than the zero baseline.

Cry vs Baseline. The analysis highlighted a right temporal cluster of two channels (19 and 22) (p<0.001) and a left fronto-temporo-parietal cluster of two channels (4 and 9) (p<0.001) where the response to the cries was significantly larger than the zero baseline (Figure II-4).

DeoxyHb.

Speech vs Cries. The cluster-based permutation analysis highlighted a right temporo-parietal cluster (channels 18, 19 and 21) where the changes in deoxygenated hemoglobin differed significantly between the two conditions. The deoxygenated hemoglobin values were lower in response to speech than in response to cries.

Speech vs Baseline. The analysis revealed a right temporal cluster of two channels (17 and 19) where the deoxy curve was significantly lower than the baseline.

Cries vs Baseline. The permutations highlighted that the deoxygenated hemoglobin values were significantly lower than the baseline in a right temporal channel (17).

3.1.7 Discussion

We carried out an fNIRS study to investigate French newborns' brain activation when they were listening to speech in their native language, and to French newborns' cries. This study was conducted to better understand what is the category "speech" for infants at birth: a broad category including speech as well as primate and human vocalizations or a specific and unique category comprising speech alone. By recording newborns' brain activity to these two types of auditory stimuli, we investigated if cries trigger similar brain activation as speech.

When directly comparing responses to speech and cries, our results show that cries trigger a significantly greater brain activation than speech in the right temporal area. The comparison with baseline shows a similar pattern: cries activated the brain more broadly than speech, triggering bilateral activation.

While in a single parietal channel (ch. 23), we also observed greater activation to speech than to cries, the hemodynamic response to speech was very brief, and had an unusual shape in this channel. This result may thus be spurious and possibly due to noise.

Cries are auditory stimuli that newborns are able to produce and use to communicate, which is not the case for speech. We hypothesize that this production advantage may explain the observed greater activation for newborn cries than for speech. This would be in line with the idea of a link between perception and production (Liberman, 1967). A few recent studies have investigated this link in young infants' brains. It has been shown that the premotor brain area gets activated when young infants were passively listening to speech (Imada et al., 2006; Kuhl et al., 2014). Moreover, infants' babbling repertoire seems to influence their speech processing and listening preference for nonwords that contain consonants that the infants themselves produce in their babbling over non-words with consonants they do not yet produce (DePaolis et al., 2011; 2013).

In our study, newborns had 1-4 days of experience with their own cries. This implies that memory formation for cry productions may be extremely rapid. It has indeed been suggested that newborns can already recognize their own cries from those of other infants (Dondi et al., 1999). Further, Polka and colleagues (2022) showed that 6- and 7-month-old infants had a listening bias for infant speech. They synthesized isolated vowels with formant values specific of an infant voice and with formants specific of a female adult voice. Using an infant-controlled looking preference paradigm, they found that young participants preferred listening to the vowels produced by an infant vocal tract, as compared to the vowels produced with adult formants. This preference increased with age between 4 and 7 months, the time during which infants enter the canonical babbling stage. To test whether this preference was simply due to the high fundamental frequency of infant voices, Polka and colleagues (2022) created infant and adult synthesized vowels that had both high fundamental frequency, and this didn't eliminate the bias for infant voices, suggesting that infants have an actual listening bias for infant speech rather than a general bias for high fundamental frequency.

To investigate the localization of the brain activations corresponding to each condition (cry/speech), we carried out analyses to identify channels that would be activated significantly more than a zero baseline by cries and by speech separately. The analysis revealed that speech in the language that the newborns heard in utero, significantly activated the left temporal area. This pattern of activation meshes well with previous results highlighting a precocious left-hemispheric activation in response to the native language, comparable to what is found in adults, in three-month-old infants (Dehaene-Lambertz et al., 2002) as well as in newborns (Peña et al., 2003; Kotilahti et al., 2009). This lateralization is the result of prenatal experience with the native language(s), as responses to unfamiliar languages are bilateral (Sato et al., 2012) or right lateralized (Vannasing et al. 2016). Our results replicate the left lateralization of the brain activation in response to the native language at birth, and thus strengthen the literature on this topic.

By contrast, cries activated a left fronto-temporo-parietal cluster that is similar to the activation found in response to speech. These shared networks could suggest a link between newborn cries and speech that trigger similar left lateralized brain activation in newborns. This could be due to the fact that both conditions are communicative: speech is how caregivers communicate with newborns, while cries represent how newborns communicate with caregivers. It has been shown that the newborn brain already responds specifically to communicative speech (Forgàcs et al., 2022) in the left fronto-temporal area. Forgàcs and colleagues (2022) compared newborns' brain activation in response to pseudowords produced by two voices speaking in turn and exchanging different pseudowords (the communicative condition), pseudowords produced by a single speaker, and two voices with one voice repeating the pseudowords that the first one just produced. They found that the left fronto-temporal area was more strongly activated by the communicative condition compared to the two other conditions. In our study, many sentences are imperative or interrogative, e.g. “Viens par là!” [Come here!]; “Est-ce que tu veux descendre?” [Do you want to get off?] or “Merci!” [Thank you!], having specific prosodies that are linked to communicative acts.

The speech condition triggered a strictly temporal activation in the left hemisphere, while the cries also triggered parietal activation. We may speculate that this reflects an activation of the premotor cortex for the cries, triggered by the listening to stimuli that newborns can themselves produce. Further studies are needed to confirm this, e.g. by comparing premotor cortex activation in newborns for stimuli that they can themselves produce vs stimuli that they are unable to generate.

We conducted the present study to better understand what counts as speech for newborns. The overlapping brain areas activated by speech and by newborn cries could be in favor of the broad category hypothesis. In the future, it will be interesting to present the same stimuli while measuring newborns' listening preference using a behavioral paradigm, similarly to previous studies (Vouloumanos & Werker, 2004; 2007; Vouloumanos et al., 2010).

Listening to newborn cries also elicited a significant brain activation in a right temporal area compared to the zero baseline, which was not the case for the speech condition. In the neuroimaging literature, right activation has been linked to prosody processing in newborns (Martinez-Alvarez et al., 2022) and young infants (Homae et al., 2006). Newborn cries can be seen as conveying both emotional prosody and, as some studies suggested (Mampe et al., 2009, Manfredi et al., 2019) the linguistic prosody of the native language the newborns heard in utero. The right-lateralized activation found in response to newborn cries could be linked to a possible processing of the cries' prosodic contour. We know that newborns are already sensitive to the prosodic features of the language they heard in utero (Abboub et al., 2016). If newborn cries already convey the typical prosodic contour of their native language, listening to native cries could trigger prosodic processing. Acoustic analyses of cries produced by newborns born in different linguistic environments are still needed to confirm or invalidate this hypothesis. Measuring the brain activation of newborns in response to cries produced by newborns from the same linguistic environment compared to cries produced by newborns from a different native language would be highly relevant in this regard. Also, a variable that should be taken into account in future studies investigating the perception of cries is whether the newborn participants have siblings. While many of the siblings' cries are likely filtered out by the womb and so do not reach the fetus, particularly loud cries may still occasionally be experienced in the womb. Indeed, if this is the case, the newborn may have been exposed to babies' or young children's crying during the prenatal period, which can influence how his/her brain reacts, at birth, to newborns cries.

Our results suggest a greater activation of the newborn brain bilaterally in response to newborn cries than in response to speech. Our study is the first to investigate how the newborn brain processes stimuli that newborns can themselves produce. This link between production and perception could explain our results. Moreover, speech and cries both activate a similar left temporal network, suggesting a link between cries and language. Is this shared activation due to the fact that both are functionally similar, i.e. they are both communicative? Is it because cries already convey the native language's prosody, to which newborns are sensitive? Or is it only due to shared acoustic features, as both speech and newborn cries are produced by the human vocal tract? To investigate these questions and better understand our results, we carried out a subsequent study by replacing cries with laughter. As with crying, laughter is auditory and communicative , produced by the human vocal tract and sharing acoustic features with speech, but it doesn't carry prosody, it is not producible by newborns themselves and it has the opposite (positive) valence.

Indeed, the valence of the stimuli may be an important factor as previous studies found that listening to peers' cries triggered signs of self-distress in newborns and can trigger crying in the listeners, a phenomenon known as "cry contagion" or "emotional contagion"

(Simner, 1971; Dondi et al., 1999; Geangu et al., 2010). We did not measure physiological signs of newborns' affective states in the current study. However, during data collection, we observed more eyebrow and mouth movements in our participants when they were listening to cries than to speech, suggesting a potential muscle activation / preparedness for crying movements. (Interestingly, this contagious crying might explain to some extent the high rejection rate, 64%, we got in the current study, as these movements created artefacts in the data.). Whether and how listening to newborn cries may impact newborns' emotional states is an important question for future studies.

3.2 Experiment *2:* Newborns' processing of baby and adult laughter

3.2.1 Introduction

In Experiment 1, we explored newborns' brain responses to speech and to newborn cries. Newborn cries are interesting stimuli for our research question about the nature of newborns' representation of speech, as they are human sounds that newborns themselves use to communicate and they are perceived as communicative by caregivers. Our results showed that newborn cries activate right parietal areas more strongly than speech. We hypothesized that this is due to the fact that cries are auditory stimuli that newborns themselves can produce. Moreover, both speech and cry significantly activated the left temporal region, which may suggest a link between cries and speech. Contrary to speech, however, cries triggered a bilateral response, as the right temporal area was also activated.

To better understand this pattern of results, we decided to run a subsequent study in which we replaced the cries by laughter. Laughter, just like cries and speech, is a communicative, auditory stimulus produced by the human vocal tract, which thus shares some acoustic features with speech (Meyer et al., 2005) and is also produced by our closest ancestors, i.e. chimpanzees (Davila-Ross et al., 2011). Speech and laughter both involve vocal and sub-glottal changes producing syllables in the case of speech and "laughter syllables" in the case of laughter. Laughter displays rich harmonic structures, with formant frequencies. However, laughter contains less fundamental frequency variation than speech, is less modified by articulation and more invariant than speech (Meyer et al., 2005; Scott et al., 2014) (Figure 3.5).

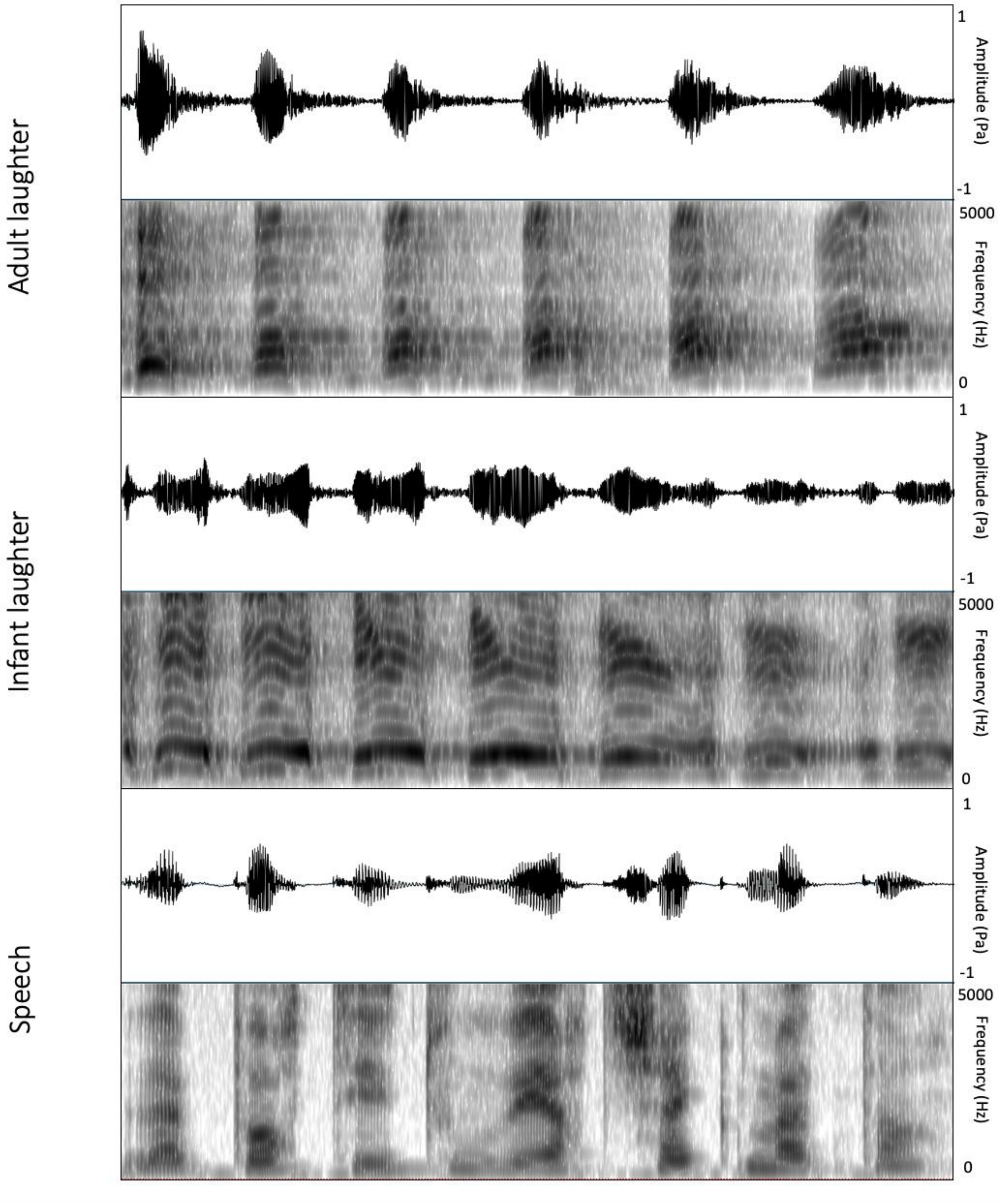


Figure 3.5 : Waveforms (upper panels) and spectrograms (lower panels) of adult laughter, infant laughter and speech examples used in this study. The time scale on the x axis is the same for all three (laughter and speech). Laughter is characterized by several subsequent bursts, corresponding to the movements of the intercostal muscles. Speech displays a greater spectral complexity than laughter.

Laughter thus shares some acoustic features with speech and are produced by the human vocal tract, like newborn cries. However, it does not have a continuous prosodic contour, while newborn cries do (Mampe et al., 2009; Manfredi et al., 2019).

Importantly, laughter has an opposite (positive) valence than cries. Comparing newborns' brain responses to auditory stimuli with positive and negative valence can inform us about the influence of valence on processing, as it has been shown in adults that it can impact the strength and localization of brain responses (Warren et al., 2006; Ethofer et al., 2007; Viinikainen et al., 2012). Further, testing sounds with different valences allows us to better explore the perimeters of newborns' category for speech, as we can assess whether and how newborns encode the affective contents of communicative stimuli.

The literature distinguishes between voluntary and involuntary laughter. It is however important to note that this distinction mainly concerns lab stimuli and that it is more questionable for laughter in daily life (Gervais & Wilson, 2005; Mazzocconi et al., 2020). Involuntary laughter is spontaneous, commonly produced in reaction to outside events, unexpected situations, etc. Voluntary laughter is a controlled production, usually associated with an intentional communicative aim. The two types of laughter have different acoustic features, spontaneous laughter being longer in duration, less voiced, higher-pitched and with higher intensity (Lavan et al., 2017). The two types of laughter are well distinguished behaviorally by adults (Bryant & Aktipis, 2014; McGettigan et al., 2015; Lavan et al., 2016) and they don't elicit the same brain activations (Szameitat et al., 2010; McGettigan et al., 2015). McGettigan and colleagues (2015) compared brain responses to spontaneous laughter and to voluntary laughter using fMRI. They found that spontaneous laughter elicited a greater BOLD response in the auditory cortex bilaterally (Heschl's gyrus and Superior Temporal Gyrus), while voluntary laughter activated the anterior medial prefrontal cortex. The stronger auditory responses could be explained by the different acoustic features of the two types of laughter, as involuntary laughter has greater intensity, while the prefrontal activation in response to controlled laughter can be explained by the listener's attempt to process its purpose and meaning (Scott et al., 2014).

Voluntary laughter has been shown to be closely linked to language and some scholars have argued that it is best understood as a linguistic unit, as it can be analyzed as part of syntax and conveys well defined meanings (Ginzburg et al., 2015; 2020). Indeed, several studies have shown that most laughs are produced during conversations, not after jokes (Provine, 2004; Vettin & Todt, 2004; Scott et al., 2014), and they carry clearly interpretable, formally analyzable meaning. For instance, compare the sentences:

(1)

a. Jill is John’s long-term friend.

b. Jill is John’s (laugh) long-term friend.

The sentence (1b) conveys the meaning that “long-term friend” is a euphemism (e.g. for “lover”, “partner” etc.).

Only a few studies addressed how the adult (Meyer et al., 2005) and the infant brain (Blasi et al., 2011; Minagawa-Kawai et al., 2011; Missana et al., 2017) processes laughter. Meyer and colleagues (2005) investigated the brain responses of adult participants, using fMRI, to speech, laughter and non-vocal sounds in the peri-auditory brain regions. Twelve adult participants were exposed to human laughs, sentential speech and non-vocal sounds while they were asked to perform a simple auditory target detection task. The laughter condition was composed of 10 female and 10 male laughing phrases, the speech condition contained short sentences produced by a female speaker and the non-vocal sound condition consisted of repetitive sounds composed of single connected frequencies. Laughter activated the right peri-sylvian areas more strongly than speech, while speech triggered greater activity in the left peri-sylvian cortex along the superior temporal gyrus and in the frontal operculum. Grouping together the responses to laughter and speech and comparing them to those of non-vocal sounds allowed the authors to investigate the overall response to human voice in general. They found that the stimuli produced by the human vocal tract elicited greater involvement of bilateral temporal and occipital sites, with a stronger involvement of the right hemisphere and especially the supratemporal sites.

What about infants’ processing of laughter? The developmental trajectory of infants’ production of smiles and laughter has been described in detail in the literature. The production of laughter appears during the first months of life, starting with the first social smiles appearing by 5-9 weeks (Wolff, 1963). The first laughter in response to physical stimulation appears at 3-4 months (Sroufe & Wunsch, 1972). Laughter-inducing events then diversify with laughter appearing in response to social games at 5 months, visual events at 7-9 months and during the creation of humorous scenes at 9-11 months (Mireault et al., 2011).

Perceptually, infants have been shown to be sensitive to emotional facial expressions already from their first few months of life and can discriminate between emotional facial expressions such as happy, angry and fearful faces (Kotsoni et al., 2001, Nelson & de Haan, 1996). Infants are also sensitive to emotions transmitted via auditory stimuli; newborns show behavioral signs of distress when listening to cries of their peers, but not to their own cries (Simner, 1971; Dondi et al., 1999). Infants also start to discriminate the emotional vocalizations of adults starting at around 5 months (Flom & Bahrick, 2007).

The sensitivity and discrimination of emotional vocalizations has also been studied at the neural level in infants (Grossmann et al., 2010; Blasi et al., 2011; Minagawa-Kawai et al., 2011; Grossmann et al., 2013; Missana et al., 2017) and in neonates (Zhang et al., 2017) using various brain imaging techniques such as ERP measures (Grossmann et al., 2013; Missana et al., 2017), fMRI (Blasi et al., 2011) and fNIRS (Grossmann et al., 2010; Minagawa-Kawai et al., 2011; Zhang et al, 2017; Maria et al., 2018 for a review). Most of these studies looked at how the infant brain processes emotions in linguistic stimuli, only a few tested responses to laughter itself (adult laughter was tested in Blasi et al., 2011; Minaga-Kawai et al., 2011; infant laughter in Missana et al., 2017). Using fMRI, Blasi and colleagues (2011) investigated the brain activation of 21 3-7-month-old infants to vocal and non-vocal sounds. The non-vocal condition was composed of environmental sounds such as toy sounds or water sounds. Three vocal conditions of different emotional valence (neutral/happy/sad) were created to investigate which brain areas were activated by nonspeech vocal emotions. The neutral condition was made of natural adult nonspeech vocalizations with no emotional content such as sneezing, throat clearing or coughing. The happy condition was composed of adult laughs and the sad condition of adults crying. Results showed no difference between the responses to the happy and the neutral conditions. However, the sad stimuli significantly activated the insula and the gyrus rectus, compared to the neutral ones. The authors suggested that their results reflect an early functional specialization for negative emotions in infants. Using NIRS, Minagawa-Kawai and colleagues (2011) recorded 4-month-old infants' brain activity to various vocal stimuli: sentences in the native language, in a non-native language, monkey calls and emotional human vocalizations with no linguistic content. The emotional condition was composed of both positive (admiration and laughing) and negative (crying and sigh) stimuli pooled together, so the two valences cannot be compared directly. However, the results showed that the emotion conditions strongly activated the right auditory area while the other conditions triggered a left-lateralized activation. This is in line with adult data showing a right-lateralized activation to vocal emotions (Meyer et al., 2005; Wildgruber et al., 2002; Wiethoff et al., 2008). In these two studies, the stimuli were vocalizations produced by adults. Interestingly, Missana and colleagues (2017) investigated infants' brain responses to vocal stimuli of different emotional valence produced by their peers, i.e. other infants. In this study, 8-months-old German infants were exposed to positive vocalizations (infant laughter), negative vocalizations (infant cries) and neutral vocalizations (adult hummed speech) while event-related brain potentials were measured. Results showed specific brain responses to infant cries, infant laughter and to emotional vocal expressions in general compared to neutral ones. More specifically, listening to infant cries triggered an early negativity in the infant brain, appearing 200ms after stimulus onset (N200), at the temporal electrodes. Infant laughter triggered a significant positivity at 300ms (P300) at central electrodes. Laughter and cries, compared to the neutral condition, also elicited a significant late positivity (between 500 and 800 ms) at central and parietal electrodes. These results highlight that 8-month-old infants are already sensitive to the vocal expressions of their peers and process negative and positive emotional vocalizations differently, both in terms of the latency and the localization of the responses.

What is the developmental trajectory of infants' processing of vocal emotions? Flom and Bahrick (2007) investigated the discrimination of emotions in 3-7-month-old infants in bimodal (auditory and visual) and unimodal situations. While 4-month-olds are able to discriminate affect in bimodal stimuli, discrimination appears only at 5 months for auditory only stimuli. To investigate how vocal emotions are perceived very early in life, Zhang and colleagues (2017) measured neonates' brain responses to different emotional prosodies using fNIRS. Neonates listened to pseudo-sentences (Chinese sentences in which the content words were replaced by meaningless words, while the function words have been kept to maintain grammatical information) produced with prosodies reflecting four different emotions: fear, anger, happiness and neutral. They found significant differences between the responses to the conditions in two channels. In a channel covering the right middle and superior temporal gyri, the three prosodies conveying emotions (fear, anger and happy) triggered significantly greater hemodynamic responses than the neutral prosody. In a second channel over the right parietal lobe and the supramarginal gyrus, fearful prosody triggered significantly greater brain activation than the happy and neutral prosodies.

Taken together, these findings show that already at birth, the newborn brain processes neutral and emotional prosodies differently. Moreover, negative emotions (i.e. fear) seem to trigger greater brain activation than positive emotions. This is in line with the results of Blasi and colleagues (2011) who found no difference in the brain responses to happy and neutral conditions, but a significantly enhanced response to sad stimuli. Similarly, Missana and colleagues (2017) found a very early negativity in response to infant cries in 8-month-old infants, reflecting a very rapid brain response to crying peers. While these studies investigated emotions in linguistic stimuli and to a lesser extent in non-linguistic vocalizations, including laughter, they did not specifically ask whether the infant brain processed laugher and other vocalizations similarly to speech or possibly as part of the same category of communicative or primate sounds.

In the present study, we, therefore, investigated how the newborn brain processes three human vocal productions: speech, infant laughter and adult laughter. The speech condition was identical to the one in Experiment 1. The infant laughter condition replaced the newborn cry condition. While infants cannot produce laughter at birth, contrary to cries, just a few months later, they become able to do so, so we hypothesized that these stimuli may still be recognized as "peer" productions and may still rely on the perception-production link hypothesized to exist for speech perception. Interestingly, this perception-production link has been observed to play a role in how adults processes laughter (Meyer et al., 2005): when adults listen to laughter, brain regions that control the motor functions of the larynx are activated. If the perception-production link also underlies infants' processing of laughter, then infant laughter may activate the newborn brain more strongly than adult laughter. We also hypothesized that if newborns process laughter similarly to adults, then they may preferentially activate the right hemisphere. Alternatively, if infants' "speech" category is broadly-based, laughter and speech may activate overlapping left-lateralized regions. We also asked whether cries (negative stimuli) elicit greater hemodynamic responses than laughter (positive stimuli) in the newborn brain, similarly to what was found in older infants. Finally, we also compared responses to the speech stimuli between Experiments 1 and 2. The speech stimuli were exactly the same in the two experiments, it would thus be feasible to predict highly similar responses. However, previous NIRS studies found different responses to the same speech stimuli in newborns as a function of the context, i.e. the other conditions presented in the study. In particular, May and colleagues (2017) conducted two separate studies with newborns, using the same paradigm, NIRS machine and settings. In a first study, newborns listened to English (native language) and Spanish (unfamiliar language) while their brain activity was recorded using fNIRS. In a following study, newborns listened to Spanish, i.e. the exact same stimuli as in the first study, and to a whistled surrogate language derived from Spanish. The brain responses to Spanish differed significantly between the first and the second study despite the stimuli being identical. Specifically, responses to Spanish were weaker when Spanish was presented together with the native language English than when it was presented together with the whistled language surrogate, as if the extent to which the newborn brain processed stimuli as language-like was relative. This suggests that the context may impact how newborns respond to speech stimuli. As we used the exact same French sentences in Experiment 1 and Experiment 2, we were able to test this context effect by statistically comparing the responses to speech in Experiments 1 and 2. Indirectly, this comparison also sheds light on the extent to which the other stimuli, i.e. cries and laughter, were processed as language-like by the newborn brain.

3.2.2 Participants

Fifteen healthy full-term newborn infants, who had been exposed at least 80% of the time to French pre- and postnatally, were included in the final analysis (11 females; mean age: 2,27 days; age range: 2-3 days, mean weight: 3240g, weight range: 2700-3795g, Apgar score 5 minutes after birth at least 8). Fifteen additional infants were tested but excluded from analysis due to poor data quality (5), fussiness or crying (9) or less than 80% exposure to French during the pre- and postnatal period (1). Recruitment was carried out following the same procedure as for the Experiment 1. All parents of the newborns tested gave written informed consent before the study, which was approved by the CPP Sud Est V ethics committee (N° ID RCB: 2018-A01136-49).

3.2.3 Stimuli

The stimuli were composed of (i) spoken French sentences produced by 10 adult female French natives (same stimuli as those used in Experiment 1), (ii) laughs produced by infants aged between 3 and 10 months (iii) laughs produced by male and female adults. Both infant and adult laughs were matched in duration with specific utterances in the speech condition (mean infant laugh length: 1111 ms, SD: 0.53, mean adult laugh length: 1110 ms, SD: 0.55, mean sentence length: 1094 ms, SD: 0.49). The duration of the stimuli did not differ significantly between the three conditions (p=0.968). The intensity of all the stimuli were normalized. However, the three types of stimuli differed significantly in their mean pitch ($p < .001$), which is inevitable given the age and gender differences between the speakers who produced them. Infant laughs have the highest mean pitch (M = 439 Hz; SD = 64), significantly higher than adult laughs (M = 317 Hz; SD = 104), which in turn have a significantly higher pitch than the spoken sentences (M = 233 Hz; SD =31). Further, while there is no significant difference in pitch excursion (difference between the maximum and the minimum pitch) between the two laughter conditions, both infant laughter (M = 247 Hz; SD = 126) and adult laughter (M = 259 Hz; SD = 156) show significantly greater pitch excursion than the speech condition (M = 197 Hz; SD = 128).

Laugh stimuli were extracted from YouTube videos. The full audio file of each video was first extracted, then the isolated laughs were segmented from the audio file using Praat. Infant laughs were elicited by tickling, visual and auditory events (parents faking a sneeze or producing weird noises, playing peek-a-boo with the baby, etc.). Adult laughs came from videos of celebrity interviews (journalists asking funny questions or making jokes) and TV series bloopers. Laughs chosen in this study are mainly involuntary, spontaneous laughs, not voluntary, conversational laughter. This is important to take into account as the two kinds of laughter trigger different brain responses, at least in adults (Szameitat et al., 2010 ; McGettigan et al. 2015).

The experiment consisted of a total of 30 blocks, 10 blocks of infant laughter, 10 blocks of adult laughter and 10 blocks of French sentences. Every block of infant laughter comprised 10 different laughs, every block of adult laughter comprised 10 different laughs and every block of speech comprised 10 different utterances, each produced by a different French speaker. Fifty-five adult laughs were produced by female voices, forty-five laughs by male voices. The adult laughs were a mix of male and female laughs for practical reasons; a large number of laughs was needed (100), which did not allow us to select only female’s laughter. We address this issue in the discussion. Each utterance, infant laugh and adult laugh was presented only once during the study. Blocks were separated by silent intervals jittered between 25 and 30 s. Within blocks, laughs and sentences were also separated by silences jittered between 500 and 1500 ms, yielding blocks of approximately 20s. The whole experiment lasted about 24 minutes (Figure 3.6).

Blocks from the three conditions were intermixed and their order was counterbalanced, such that no more than two blocks of the same condition (infant laughter/adult laughter/speech) occurred consecutively. Each newborn listened to the stimuli in a different order. Stimulus presentation within blocks was also randomized across participants.

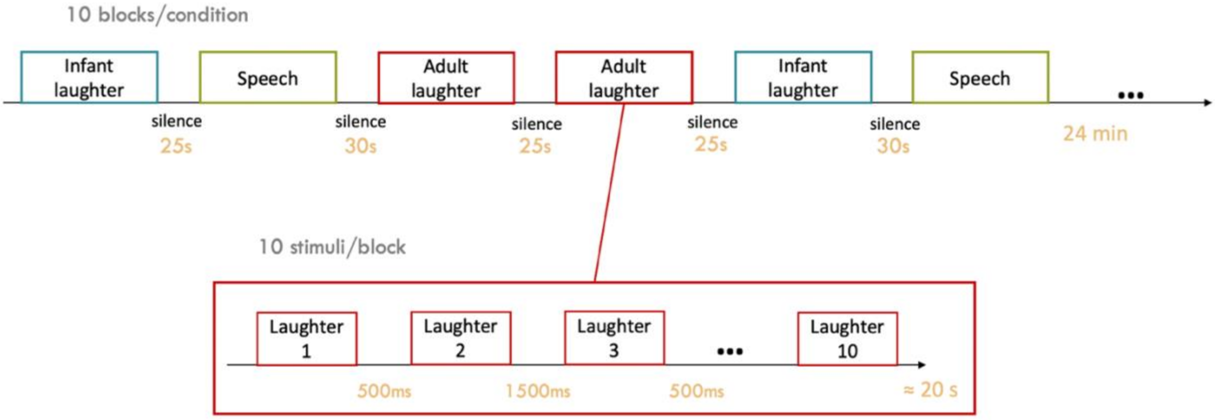


Figure 3.6 : Design of the study. The upper boxcar shows the organization and timescale of the entire experiment; the lower boxcar illustrates the organization of one single block.

3.2.4 Procedure

Same as in Experiment 1.

3.2.5 Data analysis and statistics

NIRS data pre-processing was identical to Experiment 1, except that, to further ensure appropriate data quality, we additionally performed a manual rejection of the remaining artifacts detected by visual inspection of individual data.

Statistical analyses were carried out over the concentration changes of oxyHb and deoxyHb. To compare the conditions directly, we conducted cluster-based permutation tests with a one-way ANOVA with Condition (Adult laughter / Infant laughter / Speech) as a within-subject factor. One thousand permutations were conducted under the null hypothesis, for both oxyHb and deoxyHb.

We also conducted cluster-based permutation analyses using a t-test between each condition and the zero baseline. Again, one thousand permutations were conducted under the null hypothesis, for both oxyHb and deoxyHb.

Furthermore, to investigate if, and if so, how, newborns' brain responses differed between cries (Experiment 1) and laughter (Experiment 2), we normalized the NIRS data to allow for between-subject comparisons (Otsuka et al. 2007; Berent et al. 2021) and conducted an ANOVA with Type of Stimulus (Newborn Cries/Infant Laughter) as a between-subjects factor over oxygenated hemoglobin concentration changes. (DeoxyHb is known to be less sensitive in infants, thus oxyHb is more often used, especially for small effects; Lloyd-Fox et al., 2010.)

We also compared responses to speech between Experiments 1 and 2 in an ANOVA with Experiment (Exp 1 / Exp 2) as a between-subjects factor over oxygenated hemoglobin concentration changes.

3.2.6 Results

Figure 3.7 represents newborns' grand average hemodynamic response in the 24 channels for oxyHb and deoxyHb concentration changes.

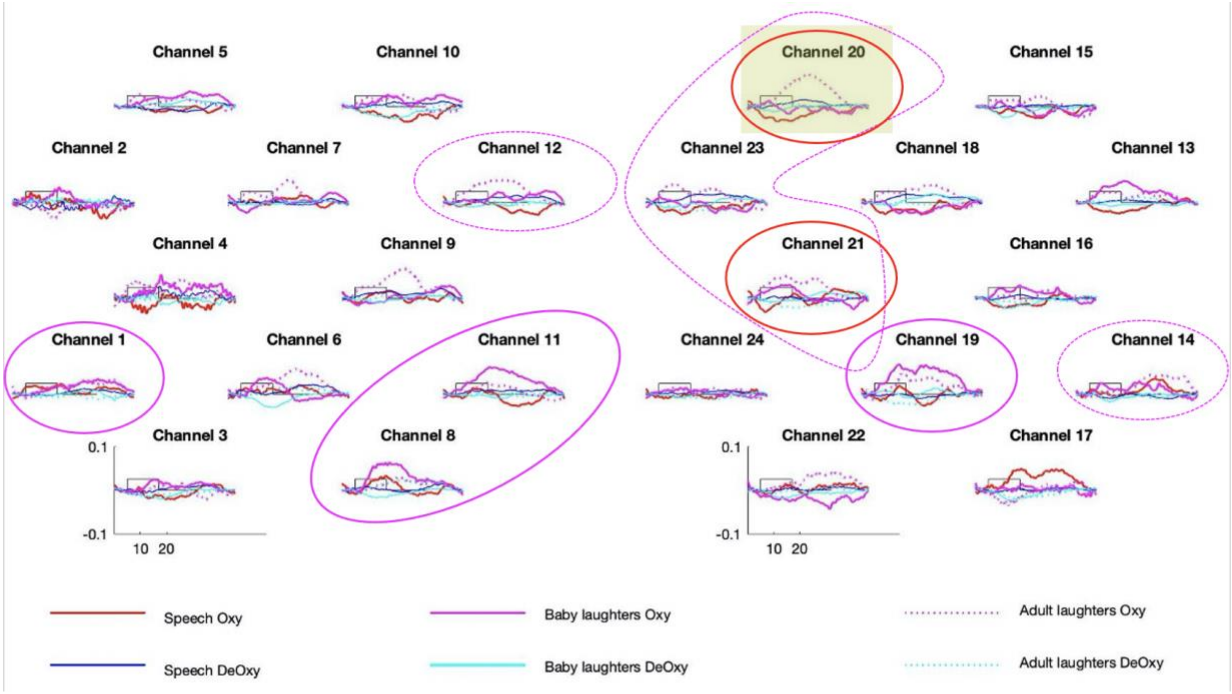


Figure 3.7 : The grand average hemodynamic responses obtained in Experiment 2. Channels are plotted following the probe placement displayed in Figure 3.3 (Panel B). The x-axis represents time in seconds; the y axis shows the concentration in mmol x mm. The rectangle along the x-axis indicates the time of

stimulation. OxyHb and deoxyHb concentration changes in response to the Speech condition are shown in red and blue, respectively (continuous lines). OxyHb and deoxyHb concentration changes in response to the Infant Laughter condition are shown in magenta and cyan, respectively (continuous lines). OxyHb and deoxyHb concentration changes in response to the Adult Laughter condition are shown in magenta and cyan, respectively (dashed lines). The channel in which the cluster-based permutations revealed a significantly different response across the three conditions is highlighted by a green rectangle. The clusters in which the Adult Laughter condition (pink dashed lines), the Infant Laughter condition (pink continuous lines) and the Speech condition (red lines) triggered a significantly different activation from the baseline are shown by colored circles.

3.2.6.1 Results of Experiment 2

OxyHb

The permutation tests comparing the conditions revealed a main effect of Condition (Adult Laughter/Infant Laughter/Speech) over oxygenated hemoglobin changes in a right parietal region (channel 20) ($p<0.001$). To understand which pairwise comparisons carried the effect, we extracted the data from channel 20 in the time window of significance highlighted by the permutation tests for the three conditions. Then, we performed an ANOVA with the within-subjects factor Condition (Adult Laughter/Infant Laughter/Speech), confirming the main effect of Condition ($F(2,14)= 4.822$, $p=0.013$). Post-hoc tests revealed that the significant difference was carried by a significantly greater activation ($p=0.016$) for Adult Laughter (M=0,06554 mmol x mm) than for Speech (M= -0,02119). The difference between Adult Laughter and Infant Laughter was marginally significant ($p=0.054$), Adult Laughter triggering greater activation than Infant Laughter (M= -0,00629).

The statistical comparisons between each condition and the baseline revealed the following.

Adult Laughter vs Baseline. The analysis highlighted a left parietal channel (channel 12) ($p<0.001$), a right frontal channel (channel 14) ($p<0.001$) and a right parietal cluster of three channels (channels 20, 21 and 23) ($p<0.001$) where Adult Laughter elicited a significantly greater brain activation than the zero baseline.

Infant Laughter vs Baseline. The cluster-based permutation analysis revealed that Infant Laughter significantly activated a left frontal channel (channel 1) ($p=.005$), a left temporal cluster of two channels (channels 8 and 11) ($p<.001$) as well as a right temporal channel (channel 19) ($p<.001$), as compared to the baseline.

Speech vs Baseline. The analysis highlighted two right parietal channels (channels 20 and 21) (p > .001), where the response to Speech differed significantly from the zero baseline, due to a negative (inverted) response.

DeoxyHb

The permutation test comparing the conditions revealed a main effect of Condition (Adult Laughter/Infant Laughter/Speech) over the deoxygenated hemoglobin changes in a right parietal region (channel 23) (p<0.001). To understand which pairwise comparisons carried the effect, we extracted the data from channel 23 in the time window of significance highlighted by the permutation test for the three conditions. Then, we performed an ANOVA with the within-subjects factor Condition (Adult Laughter/Infant Laughter/Speech). The ANOVA confirmed the main effect of Condition (p=0.019). Post-hoc tests revealed that the significant difference was explained by a significantly greater decrease of deoxygenated hemoglobin (p=0.015) for Adult Laughter (M= -0,01382 mmol x mm) than for Speech (M= 0,01163).

The statistical comparisons between each condition and the baseline revealed the following.
Adult Laughter vs Baseline. The permutations revealed a left temporal channel (channel 3) (p = 0.019) as well as a right temporal (channel 17) (p<.001) and a right parietal (channel 23) (p = 0.012) area where the response differed significantly from the baseline.
Infant Laughter vs Baseline. The permutations yielded a right parietal channel (channel 18) (p<.001) where the response was greater than the zero baseline (inverted response), and a left temporal cluster (channels 3 and 6) (p<.001) where the response significantly more negative than the baseline.
Speech vs Baseline. The analysis revealed a left temporal cluster (channels 6 and 9) (p<.001) where the response to speech was significantly greater than the baseline (inverted response).

3.2.6.2 Comparing Experiments 1 and 2: cries and laughter

To compare newborns' brain responses to newborn cries and infant laughter, we conducted a direct statistical comparison on normalized oxyhemoglobin concentration changes in the significant clusters and time windows obtained through the permutation tests in the two studies. In the cries study, the cluster-based permutation analysis conducted between the cries condition and the zero baseline highlighted a right temporal cluster comprising channels 19 and 22 and a left fronto-temporo-parietal cluster comprising channels 4 and 9. In the laughter study, a left frontal channel (channel 1), a left temporal cluster comprising channels 8 and 11 and a right temporal channel (channel 19) were significantly activated by the infant laughter compared to baseline. To conduct a direct comparison, we extracted the data from the union of the significant clusters in the two studies: a left fronto-temporo-parietal cluster (channels 1,4,8,9,11) and a right temporal cluster (19,22)

The ANOVA over oxyHb concentration changes with Stimulus (Newborn Cries/Infant Laughter) as a between-subjects factor and Hemisphere (Left/Right) as a within-subjects factor did not yield significant effects or interactions.

3.2.6.3 Comparing Experiments 1 and 2: speech

We conducted a direct statistical comparison on normalized oxyhemoglobin concentration changes between the responses to speech in Experiments 1 and 2. We extracted the data from the union of the clusters where the response to speech compared to baseline was significant within the time windows of significance. This leads to two regions of interest (ROI) comprising the left temporal channels 3 and 6 (Exp 1) and the right parietal channels 20 and 21 (Exp 2).

The ANOVA over oxyHb concentration changes with Context (Cries/Laughter) as a between-subjects factor and ROI (Left Temporal/Right Parietal) as a within-subjects factor yielded a significant main effect of Context (p=0.005), as speech triggered positive responses in the cries study (M = 0,02612 mmol x mm), but negative, inverted, responses in the laughter study (M = -0,01006) (Figure 3.8). The factor ROI was not significant, neither was the interaction Context X ROI.

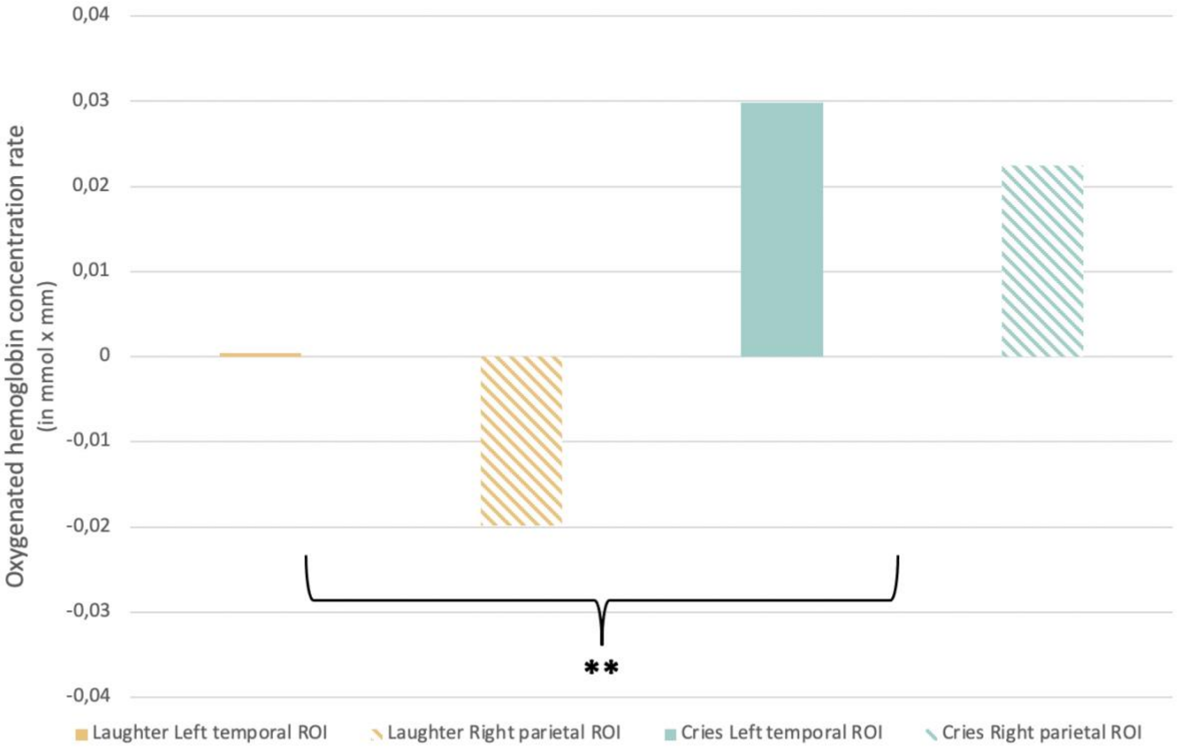


Figure 3.8 : Mean hemodynamic responses to speech in Experiments 1 (Cries) and 2 (Laughter) (in mmol x mm) in the left temporal and right parietal ROI selected for the analysis.

3.2.7 Discussion

In this fNIRS study, we measured newborns' brain responses to speech (i.e, their native language) and to adult and infant laughter. This study was a follow-up of Experiment 1, in which we measured newborns' brain responses to speech and newborns cries. In Experiment 2, we sought to investigate the role of the valence (positive vs negative) of the stimuli and the role of prosodic contours, conveyed by newborn cries (Mampe et al., 2009; Manfredi et al., 2019), but not by laughter. Further, to test infants' potential listening bias for stimuli that resemble what they themselves can produce (Polka et al., 2022), we added an adult laughter condition to compare with infant laughter.

The statistical analyses comparing responses to the three conditions revealed that adult laughter triggered significantly greater brain activation than speech in the right parietal area. Our speech stimuli were French sentences produced by female native speakers without any particular emotional valence, in a neutral way. This finding converges with previous results showing that the newborn brain displays greater hemodynamic response to emotional than to neutral prosody (Zhang et al., 2017). Moreover, several studies have shown that infants prefer to listen to a positive than a negative or neutral vocalization (Papousek et al., 1990; Kitamura & Burnham, 1998; Singh et al., 2002). Our results suggest that infants' sensitivity to positive versus neutral adult vocalizations may already be present at birth. This may be related to newborns' prenatal experience with neutral and positive vocal productions. As laughter is partly produced by rapid intercostal movements, the fetus may have a specific experience with adult laughter: when the mother laughs, the sounds heard are accompanied with movements of the mother's body that could render these stimuli particularly salient for the fetus. Prenatal experience with vocal emotions has been shown to influence newborns' perception of emotional vocalizations (Mastropieri & Turkewicz, 1999). English and Spanish newborns' eye opening and mouthing movements were measured, while they were listening to sentences produced with four different emotional affect (happy/sad/angry/neutral). Four groups of infants were tested. In two groups, the newborns listened to sentences in their native language: English newborns listening to English sentences, Spanish newborns listening to Spanish sentences. In two other groups, the newborns listened to sentences in the unfamiliar language: English newborns listening to Spanish sentences and Spanish newborns listening to English sentences. They found that newborns display different responses to the emotional conditions only when the sentences were produced in their native language, highlighting the influence of prenatal experience with maternal emotional prosody. Interestingly, when newborns listened to the sentences in their native language, the highest score for eye opening was found for happy speech, as compared to sad or neutral speech. Moreover, when newborns had their eyes closed at the beginning of the stimulation, they tended to keep their eyes closed throughout the study, except in the happy condition, during which they opened their eyes. This precocious attraction for happy adult speech at birth could explain the greater hemodynamic response we found in our experiment for adult laughter compared to neutral adult speech in the newborn brain. We may hypothesize that this precocious attraction for happy adult speech is a manifestation of an innate self-preservation instinct shaped by evolution: a caregiver's happy, positive attitude is a sign of a more attentive, reassuring, inclined-to-care environment. Adult laughter may thus be more attractive and more relevant for infants than infant laughter, which might explain why adult laughter triggered slightly greater brain activation than infant laughter.

Furthermore, adult laughter activated the bilateral parietal regions, with a larger involvement of the right hemisphere (three right channels vs one single left parietal channel), as well as the right frontal area. This right-dominance is in line with adult data showing right-lateralized processing of spontaneous laughter (Meyer et al., 2005). Our adult laughs stimuli were a mix of male and female laughs, while the speech stimuli were produced only by female voices. It would be interesting to test the effect of gender by presenting only male or only female laughs. However, as laughter does not have proper pitch, the acoustic difference between male and female laughs should be less important than between male and female speech.

Infant laughter, by contrast, triggered bilateral temporal and left frontal activation. This pattern of activation is similar to what we found in Experiment 1 in response to newborn cries. Newborn cries elicited significant activation in the right temporal and left temporo-frontal areas. Indeed, the direct statistical comparison between the responses to infant laughter and newborn cries revealed no significant differences in the bilateral temporal region. We may infer that infants' emotional vocalizations, whether positive or negative, activate a common, shared bilateral temporal and left frontal network in the newborn brain. While previous studies suggest that sad adult vocalizations may activate the infant brain more than happy adult vocalizations (Blasi et al., 2011; Zhang et al., 2017), positive and negative emotional vocalizations produced by peers (i.e., infants) both triggered specific brain responses in 8-month-old infants (Missana et al., 2017). This is in line with our results, which did not show greater activation for negative than positive infant vocalizations in newborns.

The speech condition triggered significant activation in the right parietal region, due to a negative (inverted) response. Indeed, speech triggered flat or inverted responses in this experiment (Figure 3.7). We thus did not find the usual left-lateralized activation for the native language found in several fNIRS studies with newborns (Experiment 1; Peña et al., 2003; Sato et al., 2011), including our own Exp 1. Indeed, even though the stimuli were identical between Experiments 1 and 2, the responses we obtained were significantly different. This pattern of results has also been found in a fNIRS study with newborns conducted by May and colleagues (2017). The authors carried out two subsequent studies. First, English-exposed neonates were presented with sentences in their native language and in an unfamiliar language (Spanish), played forward and backward (English Forward/English Backward/Spanish Forward/Spanish Backward). They found a forward advantage for the native language, but not for Spanish. In a following study, a new set of twenty English learning-neonates were exposed to the same two Spanish conditions compared to forward and backward Silbo Gomero, an unfamiliar surrogate whistled language. In this case, forward and backward Spanish elicited significantly different responses, which was not the case when Spanish was presented in comparison with the native language, despite the fact that the exact same Spanish stimuli were used in the two studies. The authors suggested that the context may influence what is processed as language by the newborn brain. Similarly, in our study the context may explain why speech gave rise to canonical hemodynamic responses in the left temporal area in Experiment 1, but to inverted, bilateral responses in Experiment 2. First, in Experiment 1, speech produced by adults speakers is presented along with newborns cries, while in Experiment 2, speech is presented along with infant and adult laughs. Secondly, speech is compared with stimuli with opposite emotional valence in the two experiments. As newborns are more sensitive to happy than to neutral adult speech (Mastropieri & Turkewicz, 1999), laughter (i.e. happy stimuli) may be more salient than speech [Experiment 2], which is not the case for cries (i.e. negative stimuli) [Experiment 1] and this may explain why the brain responded differently to the same exact speech stimuli in the two experiments. Additionally, the acoustic proximity between stimuli may also play a role. Speech may elicit canonical responses when compared to cries but not when compared to laughter, because cries and speech both have prosodic contours, while laughter differs more from speech acoustically.

3.3 General discussion

In this chapter, we investigated newborns' brain responses to speech and to non-speech human vocal sounds, specifically newborn cries (Experiment 1), infant laughter and adult laughter (Experiment 2). Newborn cries elicited greater hemodynamic response than speech in the right temporal area, while adult laughter triggered greater responses than speech in the right parietal area. The non-speech human vocal sounds used in these two experiments had an emotional valence, while the French sentences used in the Speech condition were neutral. This might explain why speech elicited less activation than the nonspeech stimuli, as newborns have been found to be sensitive to emotional vocal productions, as compared to neutral ones (Mastropieri & Turkewicz, 1999; Zhang et al., 2017). Future studies could test whether using speech stimuli with positive/negative emotional valence as a comparison for cries or laughter may lead to different results.

The aim of the two studies was to investigate the nature of the category "speech" at birth. Existing behavioral results suggest that newborns may display a preference for a broad category including speech as well as other primates vocalizations (Vouloumanos et al., 2010). Our results, converging with previous studies (Peña et al. 2003, Sato et al., 2011), show that speech triggers a specific brain specialization already at birth. In Experiment 1, speech elicited a significant left-lateralized activation in the fronto-temporal areas. Interestingly, however, this was modulated by the context, as speech presented in the context of laughter gave rise to an inverted hemodynamic response. Importantly for our research question, both of these patterns differed from the responses elicited by cries and laughter: these non-speech stimuli both triggered bilateral activation. This was independent of whether the vocalizations were such that newborns could produce them or not.

Taken together, our results thus suggest that speech in the native language already triggers specialized brain responses at birth, different from non-linguistic, but communicative sounds such as cries and laughter. Newborns' representation of speech is thus already specific enough not to include these categories of sounds, despite the fact that they are acoustically similar, communicative and can be produced by the infants themselves.

For newborns and infants, recognizing speech and treating it as a special stimulus is fundamental for language acquisition. Experiments 1 and 2 have uncovered a specific brain activation for speech already at birth. To better understand how this specialization leads to language learning, we will investigate, in Chapter 4, if speech already elicits specific learning mechanisms compared to non-linguistic stimuli at birth.

4 Does rule learning apply to non-speech sounds in the newborn brain? (Experiment 3)

This chapter is based on the publication Nallet et al. (2023).

4.1 Introduction

At birth, newborns already show remarkable speech perception abilities. They recognize and prefer speech over similarly complex sine wave analogs (Vouloumanos & Werker, 2004), prefer their mother's voice compared to unfamiliar female voices (DeCasper & Fifer, 1980; Moon, 2017) and are able to discriminate their native language from an unfamiliar language (Mehler et al., 1988; Moon et al., 1993). These sophisticated speech perception abilities pave the way for the discovery of their native language.

One of these early abilities is newborns' capacity to detect and learn structural regularities, in particular repetition- or reduplication-based structures, from the speech input (for an exhaustive review, see de la Cruz-Pavía & Gervain, 2021). When neonates listen to repetition-based sequences of syllables of the form AAB (e.g. "babamu") and random sequences of the form ABC (e.g. "bamuge"), they show greater hemodynamic responses to the AAB sequences in the bilateral temporal and left inferior frontal regions (including Broca's area) (Gervain et al., 2012) than to the random ABC sequences. This pattern of responses suggests that newborns can discriminate and differently process repetition-based and random, diversity-based structures. Newborns respond similarly to repetitions in sequence-final positions (ABB, e.g. "mubaba"; Gervain et al., 2008).

What is the nature of the mechanism that allows newborns to represent such structures? The empirical results suggest that two different mechanisms may be at work (Gervain et al., 2008; de la Cruz-Pavía & Gervain, 2021). Greater brain responses to the repetition-based than to the diversity-based structures have been observed in the left frontal areas from the beginning of the experiment (Gervain et al., 2008), suggesting an immediate and automatic recognition of repetition (Endress et al., 2009). Additionally, the differential response increased over the time course of the study, suggesting a possible mechanism responsible for the extraction of the abstract underlying reduplication regularity that is common across all ABB sequences. Indeed, as all sequences were unique, never repeated during the study and were made up of a relatively large inventory of different syllables, the only commonality across them was the underlying, abstract ABB structure. Thus, continued increase in response to ABB sequences over the course of the study can only be attributed to the extraction of this common underlying pattern across them. This interpretation is supported by the left frontal location of the effect (Gervain et al., 2008).

Newborns thus show basic abilities to learn structural regularities in speech. Is this sensitivity to repetition-based structures specific to language? In the visual domain, it has been shown that 3-4 month-old (Ferguson et al., 2018) and 7-month-old (Saffran et al., 2007) infants can discriminate ABB and ABA patterns behaviorally when presented with pictures of animals. In the auditory domain, Dawson and Gerken (2009) demonstrated that 4-month-old infants can learn repetition-based regularities carried by musical chords and tones. These results suggest that the ability to extract repetition-based regularities is not restricted to speech in infants older than newborns at the behavioral level. However, the question remains whether this is also the case for newborns.

Is this ability specific to speech at birth? Or can newborns also discriminate repetition-based sequences from diversity-based ones when those are implemented with musical tones, i.e. non-linguistic auditory stimuli? If so, are the underlying brain mechanisms similar or do they differ between linguistic and non-linguistic auditory stimuli? Specifically, is there a domain-general mechanism allowing the extraction of repetition from both linguistic and non-linguistic stimuli, or does the extraction of repetition rely on different mechanisms in the linguistic and non-linguistic domains? To address these questions, we tested whether newborns are sensitive to repetition-based regularities when those are implemented with musical tones. We followed the same experimental design as Gervain and colleagues (2008, 2012), comparing infants' responses to blocks of AAB and ABC sequences, except that syllables were replaced by tones. Newborns' brain activity was recorded using functional Near-Infrared Spectroscopy (fNIRS), as in the previous studies.

Differential responses to AAB vs. ABC sequences indicate newborns' ability to discriminate repetition-based sequences from diversity-based ones. As an advantage of using a brain imaging technique, we can go beyond this binary response to test not only the presence or absence of discrimination, but also whether the mechanisms underlying the processing of these sequences in speech and in non-speech stimuli are shared or different. It may be the case that while newborns can discriminate AAB and ABC sequences both in speech and tones, they process the two types of stimuli differently.

Previous results suggest that stimulus complexity and variability play a role in the observed responses. Indeed, Bouchon and colleagues (2015) replicated Gervain et al.'s (2012) study, but used a much smaller phoneme and syllable repertoire. With such reduced stimulus variability, the authors still found greater activation in response to the ABB than to ABC sequences. But while, in Gervain et al.'s (2012) study, responses to ABB sequences increased over time, in Bouchon et al.'s (2015) experiment, responses to the ABB structure remained stable over the course of the study, whereas responses to ABC increased. These results offer two critical lessons. First, they suggest that the variability and the complexity of the stimuli play a role in the temporal dynamics of the response patterns found, as neural responses show different degrees of habituation to the repeated presentation of stimuli as a function of stimulus complexity. Second, these results demonstrate that the temporal dynamics of neural responses can provide relevant insights into the neural mechanisms that support the detection of repetition. Accordingly, in the present study, we both compare the time-averaged responses and explore the temporal dynamics of infants' responses to ABB and ABC sequences in tones.

4.2 Participants

Thirty-two healthy, full-term newborns (Apgar score >= 8) participated in the study. Of these, data from 23 newborns (mean age: 1,18 days, range: 0-3 days; 12 females) were kept for the final analyses. Nine infants were excluded for crying and fussiness (n=6), parental intervention (n=1) or poor data quality (n=2). Rejection for data quality was performed in batch prior to statistical analysis. According to parental report, the infants did not receive greater than average prenatal exposure to music (specifically, mothers reported listening to music only occasionally), and the parents were not musicians. The infants' parents or legal guardians provided written informed consent prior to participation. The study was approved by the ethics boards of the University of British Columbia and BC Women's Hospital.

4.3 Stimuli

Two artificial grammars were used in the study: AAB and ABC (Figure III-1). The AAB grammar generated sequences of three musical tones, whereby the first and second tones were identical, whereas the ABC grammar generated sequences with three different tones. The sequences were identical to those used in Gervain et al. (2012) except that they were implemented with musical tones rather than syllables. Each syllable of the original AAB and ABC sequences in Gervain et al. (2012) was mapped onto a piano tone for a total of 20 tones corresponding to the 20 syllables used in the Gervain et al. (2012) study. The tone repertoire included C, D, E, F, G, A, B, c, d, e, f, g, a, b, c', d', e', f', g', a' (reported here using ABC notation https://abcnotation.com, C denotes middle C). Thus, the distribution and frequency of occurrence of syllables and tones were fully matched between the two studies. Furthermore, the AAB tone grammar was matched to the ABC tone grammar on (1) the overall frequency of each tone, (2) the frequency of each tone in each sequential position, and (3) all acoustic characteristics. Additionally, the distribution of transitional probabilities (TPs) between adjacent tones was equated by keeping the TPs as high between certain designated AB tone pairs, as they were between the repeated AA tones. Furthermore, when mapping the syllables to tones, the choice of tones was optimized so as to avoid dissonant tone sequences. Tones were generated using the flabc software (https://flabc.software.informer.com) with the acoustic grand piano as the MIDI instrument. Tones lasted 510 msec (i.e., the default quarter note setting in flabc), sequences thus lasted 1530 msec.

The design of the study (Figure 4.1) was identical to the one used in Gervain et al. (2012). Thus, the sequences were presented in simple blocks consisting of 10 tone sequences from a given condition (AAB or ABC) for a total of 14 blocks per condition. All 140 tone sequences for a condition were unique. No sequence was ever repeated. Within blocks, sequences were separated by pauses of varying duration (0.5–1.5 sec), yielding blocks of approx. 25 sec (Figure 4.1). Blocks were also spaced by silence periods of varying duration (25–35 sec) to avoid inducing phase-locked brain responses. The 28 blocks were presented in an interleaved fashion. The order of blocks was pseudorandomized and counterbalanced across participants. At most two consecutive blocks of the same condition were allowed.

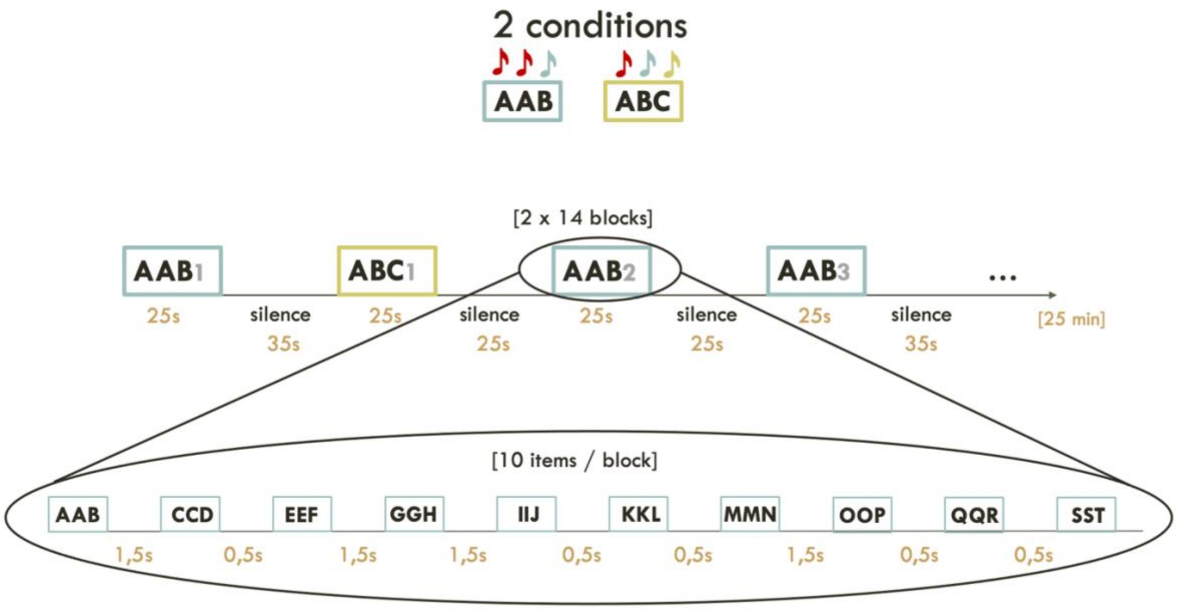


Figure 4.1 : Design of the study. The upper boxcar shows the timescale of the entire experiment; the lower boxcar illustrates the organization of a block.

4.4 Procedure

We used the same procedure as in Gervain et al. (2012). Infants were tested with a HITACHI ETG-4000 NIRS machine (source detector separation: 3 cm; two continuous wavelengths of 695 and 830 nm) in a dimly lit, sound-attenuated room at BC Women's Hospital, lying in their cribs throughout the test session. At least one parent was present at all times. Babies were tested while asleep. Indeed, newborns that are awake or awaken during the study often produce too many movement artifacts or start crying / fussing so they rarely contribute data or finish the study. Moreover, sleep states in newborns are very different from those in older infants or adults – so much so that the sensitivity of the auditory cortex to sounds is quite similar during sleep and wakefulness in newborns (Fifer et al. 2010), which is not the case later in life. The NIRS optical probes were placed on infants' heads bilaterally (12 channels per hemisphere; Figure 3.3) using the tragus, the vertex, and the ears as surface landmarks (Gervain et al., 2008; Peña et al., 2003). Sound stimuli were administered at an intensity of 60dB/SPL through two loudspeakers positioned at a distance of 1.5 m from the babies' heads, at an angle of 30°, and elevated to the same height as the infants' cribs. A portable Macintosh computer played the stimuli and operated the NIRS machine, running the PsyScope experimental software (http://psy.ck.sissa.it). The NIRS machine used 0.7 mW laser power. [The stimuli and the design of this study were created by Judit Gervain, who also collected the newborn data.]

4.5 Data analysis and statistics

Changes in the concentration of oxygenated hemoglobin (oxyHb) and deoxygenated hemoglobin (deoxyHb) were calculated from the absorption of near-infrared light as metabolic indicators of neural activity. Data were band-pass filtered between 0.01 and 0.5 Hz. Movement artifacts, defined as concentration changes larger than 0.1 mmol x mm over 0.2 msec, were removed by rejecting block-channel pairs where artifacts occurred. For the nonrejected blocks, a baseline was linearly fitted between the means of the 5 sec preceding the onset of the block and the 5 sec starting 40 sec after the onset of the block (25 sec of stimulation plus 15 sec of resting period). To further ensure appropriate data quality, we then performed a manual rejection of the remaining artifacts. Newborns were included in the analysis if at least 40% of their data was found valid after pre-processing.

Statistical analyses were carried out over the concentration changes of oxyHb and deoxyHb. First, we conducted a cluster-based permutation analysis (Maris & Oostenveld, 2007) in order to identify clusters of channels in which a significant difference in concentration changes between the two conditions were found. This analysis can define regions of interest (ROIs) and time windows of interest in a non-arbitrary and data-driven manner, while safeguarding against multiple comparisons (Benavides-Varela & Gervain, 2017; Abboub et al., 2016, Mahmoudzadeh et al., 2013). The definition of spatial clusters was anatomically constrained, as determined by a localization analysis (Abboub et al., 2016; Figure 2B). One thousand permutations were conducted under the null hypothesis, for both oxyHb and deoxyHb.

We then performed a linear mixed effects model over concentration changes in the clusters and time windows identified by the permutation analysis, with Participants as a random factor and Condition (AAB / ABC), ROI (Left Temporal cluster / Left Fronto-Parietal cluster / Right Temporal cluster / Right Fronto-Parietal cluster) and Time Course (Blocks 1-14) as fixed factors in order to assess differences across conditions in a temporally dynamical fashion. We tested all the possible models. Model selection was based on the Akaike Information Criterion (AIC) (Claeskens and Hjort, 2008).

Furthermore, to directly compare newborns' responses to the regularities in tones and speech, we normalized the NIRS data to allow for between-subject comparisons (Otsuka et al. 2007; Berent et al. 2021) and conducted an ANOVA with Type of Stimulus (Tones/Syllables) as a between-subjects factor and Regions of Interest (ROI; frontal/temporal), Hemisphere (Left Hemisphere, LH/Right Hemisphere, RH) and Condition (Repetition/Non-Repetition) as within-subjects factors.

4.6 Results

4.6.1 Discrimination of AAB and ABC structures with tones

OxyHb.

The grand average results for oxyHb and deoxyHb concentration changes are presented in Figure 4.2. The cluster-based permutation analysis revealed a fronto-temporal and a fronto-parietal cluster in each hemisphere in which brain activity significantly differed between the AAB and the ABC conditions (Figure 3). In the left hemisphere (LH), the fronto-temporal cluster included channels 1 and 3 ($p = 0.002$), the fronto-parietal cluster channels 5, 7, 9 and 12 ($p < 0.001$). In the RH, the fronto-temporal cluster included channels 17, 22 and 24 ($p < 0.001$), while the fronto-parietal cluster included channels 13, 15, 16 and 20 ($p < 0.001$).

DeoxyHb.

The cluster-based permutation analysis over the concentration changes of deoxyHb revealed that brain activity significantly differed between the AAB and ABC conditions in channel 12 in the left parietal region ($p < 0.001$), in channel 23 in the right parietal region ($p < 0.001$), in channels 15 and 18 in the left fronto-parietal region ($p < 0.001$) and in channels 14, 16, 17, 19, and 22 in the right fronto-temporal region ($p < 0.001$). The results obtained for deoxyHb were thus largely congruent with those found for oxyHb.

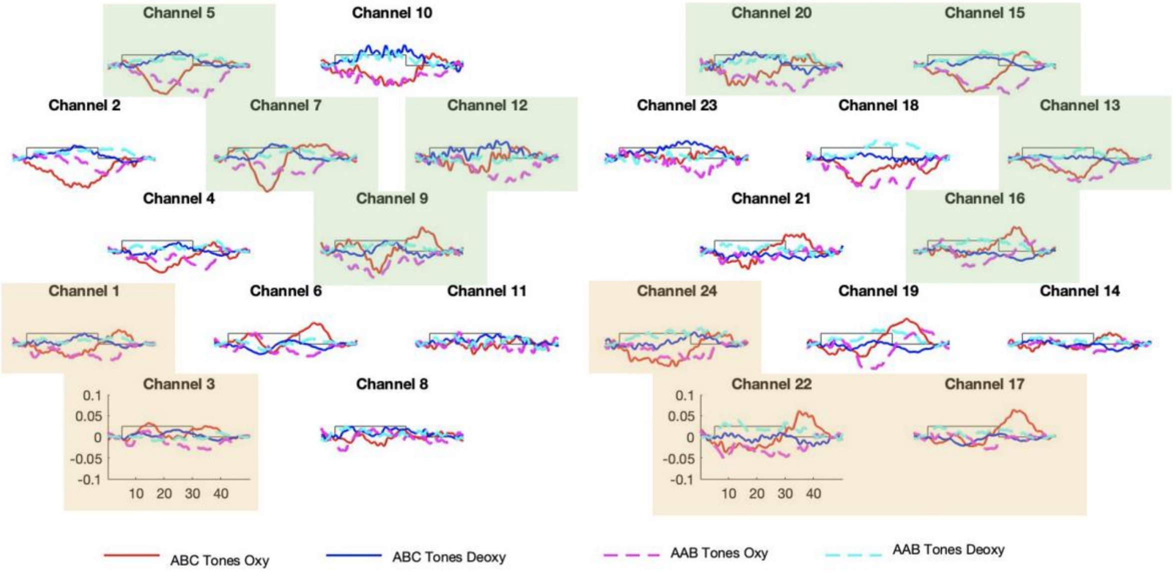


Figure 4.2 : The grand average hemodynamic response. Channels are plotted following the probe placement displayed in Figure 3.3. The x-axis represents time in seconds; the y axis shows the concentration in mmol x mm. The rectangle along the x-axis indicates the time of stimulation. OxyHb and deoxyHb concentration changes in response to the ABC condition are shown in red and blue, respectively (continuous line). OxyHb and deoxyHb concentration changes in response to the AAB condition are shown in magenta and cyan, respectively (dashed line). The four significant clusters revealed by the cluster-based permutation analysis over oxyHb concentrations are overshadowed.

The best fitting linear mixed effects model carried out on oxygenated hemoglobin concentration changes was the one with Time Course and Condition as fixed factors. This model revealed a significant main effect of Time Course ($p < 0.001$) as well as a significant main effect of Condition ($p < 0.001$). The main effect of Condition was due to a larger, i.e., more negative, inverted response to the AAB than to the ABC sequences. The best fitting linear mixed effects model when considering deoxygenated hemoglobin concentration changes was the one with Condition as a fixed factor. This model highlighted a significant main effect of Condition ($p = 0.015$). This pattern of results confirms that newborns can discriminate repetition-based sequences of musical tones from random ones. Additionally, as the main effect of Time Course indicates, responses changed significantly over the time course of the study. This was due to a general decrease in response amplitude over time for both conditions (Figure 4) reflecting neural habituation. As post hoc analyses to explore this main effect, we ran correlations between Time Course and Concentrations in each ROI. The correlation was significant in the left fronto-temporal cluster (channels 1 and 3) for the

ABC condition (r = -0.605, p = 0.022) and in the right fronto-temporal cluster (channels 17, 22 and 24) for the ABC condition (r = -0.728, p = 0.003) as well as for the AAB condition (r = -0,626, p = 0.017).

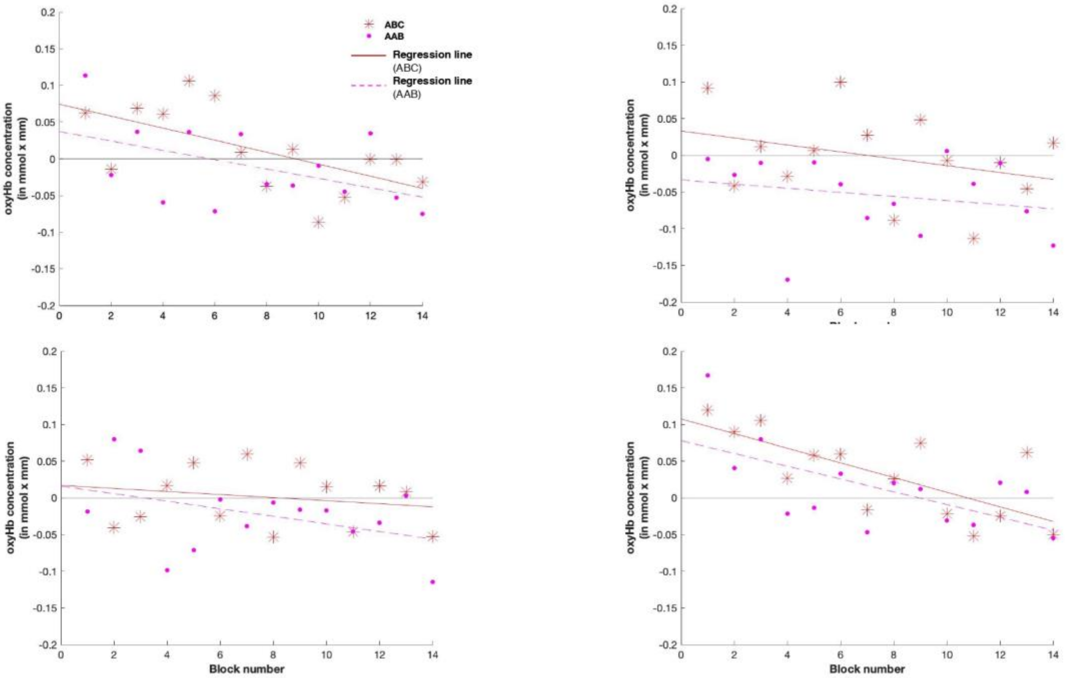


Figure 4.3 : Changes in the concentration of oxygenated hemoglobin over the time course of the study for each block in the four clusters: (A) left fronto-temporal cluster comprising channels 1 and 3, (B) left fronto-parietal cluster comprising channels 5, 7, 9 and 12, (C) right fronto-parietal cluster comprising channels 13, 15, 16 and 20, (D) right fronto-temporal cluster comprising channels 17, 22 and 24.

4.6.2 Comparing tones and speech

To compare newborns' brain responses to the regularities implemented by tones and by speech, we conducted a direct statistical comparison on normalized oxyhemoglobin concentration changes in the significant clusters and time windows obtained through the permutation tests in the two studies. For the tones study, we used the cluster of channels reported above (Figure 4.3), i.e. channels 1 and 3 for the left temporal region, channels 5, 7, 9 and 12 for the left frontal region, channels 17, 22 and 24 for the right temporal region and channels 13, 15, 16 and 20 for the right frontal region. For the speech study (Gervain et al., 2012), we performed the same cluster-based permutation analysis as for the tones study to identify significant clusters. This analysis yielded a left temporal cluster composed of channels 3 and 6, a left frontal cluster composed of channels 4 and 9, a right temporal cluster including channels 16 and 19 and a right frontal cluster including channel 20.

We then performed an ANOVA over oxyHb concentration changes with Type of Stimulus (Tones/Syllables) as a between-subjects factor and Region of Interest (ROI; Temporal/Frontal), Hemisphere (LH/RH) and Condition (Repetition/Non-Repetition) as within-subjects factors. The main effect of Type of Stimulus was significant (F(1.32)=5.905, *p*=0.021), as speech triggered a larger, positive, response (M = 0,0176 mmol x mm), while tones triggered a negative, inverted response (M = -0,0056 mmol x mm). The interaction Condition X Type of Stimulus was also significant (F(1.32)=9.071, p=0.005). Post-hoc comparisons revealed that this was due to significantly more negative responses to repetitions (M = -0,0204 mmol x mm) than to non-repetitions (M = 0,0092 mmol x mm) for tones, as well as to greater, positive response for speech (M = 0,0264 mmol x mm) as compared to negative, inverted response for the tones (M = -0,0204 mmol x mm). The main effect of ROI was also marginally significant (F(1.32)=3.861, p=0.058), due to larger brain responses in the temporal (M = 0,0112 mmol x mm) than in the frontal ROIs overall (M = 0,0009 mmol x mm). The other factors and interactions were not significant.

4.7 Discussion

We conducted an fNIRS study to investigate whether newborns can detect and learn a repetition-based regularity presented with tones, as they do with syllables. Newborns show greater brain activations in response to trisyllabic pseudo-words containing an initial (e.g. “babamu”) or a final repetition (e.g. “mubaba”) compared to trisyllabic random pseudowords (e.g. “mubage”) (Gervain et al. 2008; 2012). By mapping each syllable onto a unique tone, we wanted to test whether this precocious ability was specific to speech or extended to other auditory stimuli. Results show significantly different patterns of brain activation between the repetition and the non-repetition conditions in bilateral temporal and fronto-parietal areas. This is due to larger inverted (negative) responses to the AAB sequences than the ABC sequences. Thus, our findings suggest that newborns’ ability to detect repetition is not specific to speech, it also applies to musical tones.

This study, while based on a relatively small group of babies, can be considered within the larger framework of NIRS studies investigating young infants’ abilities to extract repetition-based regularities from language and other stimuli. To statistically assess the robustness and replicability of these results and to identify moderating factors that may impact the effects, we have recently conducted a meta-analysis of all published and unpublished NIRS studies within this paradigm (Gemignani et al., 2023), and found a consistent, replicable, low-to-medium sized effect for repetition-based rule learning in newborns from speech stimuli with a sample size of 355 newborns from 23 studies. This meta-analysis thus confirms the robustness of the repetition-based rule-learning effect and provides converging evidence with the current results.

However, the patterns of brain activation we observed to tones differed from that observed with spoken stimuli. The first difference is in the localization of the responses, as suggested by the distribution of the significant clusters. Repetitions in syllables trigger greater differential activation in the temporal and frontal areas bilaterally, but with a greater involvement of the left hemisphere (Gervain et al. 2008, 2012). By contrast, brain responses to the tone sequences were not left lateralized, but rather symmetrical between the two hemispheres. Larger activation in the left hemisphere in response to linguistic regularities is consistent with the general left hemispheric specialization of the brain for language, which is well documented in adults (Zatorre & Belin, 2001; Hickok and Poeppel 2007) as well as in infants (Dehaene-Lambertz et al. 2002; Peña et al. 2003; Minagawa-Kawai et al. 2011). By contrast, music has often been observed to activate the right hemisphere (Peretz & Zatorre, 2005). Kotilahti and colleagues (2009) compared brain activation to language (sentences from a fairytale read in Finnish by a Finnish actress) versus to music (excerpts from a Mozart piano concerto) and they found a left lateralization of the brain response only for the language condition .Whether our stimuli may be considered truly musical is an open question, as the assignment of tones to structure was determined randomly, but they were made up of piano tones, which was sufficient to trigger a non-left-lateralized response.

A second difference between the responses to regularities implemented with speech and tones lies in the nature of the hemodynamic response, as shown qualitatively, by the shape of the hemodynamic responses, and quantitatively, by the direct statistical comparison of the two studies. For syllables, newborns showed canonical (positive) responses, while in our study, responses were inverted (negative). How to best interpret inverted hemodynamic responses in infants is still under debate (Issard & Gervain 2018). Such responses have been related to a number of physiological factors, such as the immaturity of infants' vasculature, brain and the neurovascular coupling, but also to experimental factors, such as deactivation due to experimental design, ill-formed, unusual or degraded stimuli etc. In the current study, it is the result of a third difference between the responses to speech and tone stimuli, namely their temporal dynamics. Specifically, responses to syllables and tones differ when we look at them over time. In Gervain et al. (2008), the results show a greater activation for the ABB sequences from the beginning of the experiment, which further increased over the time course of the study. By contrast, in our study, responses to the tone sequences decreased over time, resulting in inverted responses after the first few blocks. This kind of deactivation over repeated stimulus presentation corresponds to a classical habituation effect.

Why did tone sequences trigger habituation, but syllables did not? The literature on infants' neural responses to repetition-based regularities offers two possible, not necessarily mutually exclusive explanations. First, Bouchon and colleagues (2015) replicated Gervain et al.'s (2008) study, but drastically reduced the variability of the speech stimuli, decreasing the number of trisyllabic sequences from 280 to 24 and increasing the frequency of occurrence of each item from 1 to 6. Contrary to the gradually increasing response to repetition sequences found in Gervain et al. (2008), i.e. a "repetition-enhancement" effect, Bouchon and colleagues (2015) observed a stable response over time. This suggests that, similarly to adults (Henson and Rugg, 2003), the habituation of infant brain to the stimuli depends on their variability/complexity. In our study, we have found decreasing activation over the time course of the study, often termed "repetition-suppression" in the fMRI literature (Henson and Rugg, 2003). Despite the that our study perfectly match the frequency of occurrence and distribution of the tones and syllables (as our tone stimuli were constructed by replacing each syllable of Gervain et al. 's (2008) with one tone), linguistic syllables are still inherently more complex than tones at the acoustic level. This lower acoustic complexity is one possible factor that may have contributed to the habituation effect and thus the inverted hemodynamic responses we observed, just as lower variability in frequency of occurrence did so in Bouchon et al. (2015). This interpretation is empirically supported by the fact that while for speech stimuli, the repetition-based regularities differentially triggered enhancement over the course of the study, for tones, both the repetition- and the diversity-based sequences gave rise to a habituation response. In other words, the deactivation was common to all tone stimuli irrespective of structure, suggesting that it is related to their general acoustics.

A second, and possibly related explanation is that linguistic stimuli, as identified acoustically by their characteristic features and complexity, may preferentially trigger the extraction of rule-learning mechanisms that are distinct from those triggered by tones. In a recent study, Berent and colleagues (2021) asked whether infants' ability to extract repetition-based regularities from language input is restricted to the spoken modality, or if it also extends to the visual modality. They thus contrasted sign-naïve infants' brain responses to repetition-based regularities carried by linguistic signs (novel syllables in American Sign Language) as well as to matched non-linguistic visual control stimuli. While infants were able to extract the repetition-based regularity from both types of stimuli, they showed opposite patterns of activations. In the signs condition, a greater activation was found for the repetition sequences than for the random ones, while the random sequences elicited greater activation than repetitions in the visual controls. Even more importantly, infants' brain responses to repetitions implemented with signs were very similar to those found for speech in Gervain et al. (2008; 2012). Despite the highly different sensory characteristics of speech and sign, repetition-based regularities elicited similar brain responses. Berent and colleagues (2021) interpreted these findings as a double dissociation between the sensory modality of stimuli (auditory or visual) and their linguistic status (linguistic or non-linguistic). Linguistic stimuli trigger similar brain responses, irrespective of their sensory properties, while non-linguistic stimuli yield a different response pattern despite perceptual similarities.

Our results are in line with this double dissociation hypothesis: despite being in the same perceptual modality as speech, tones trigger different brain responses. Taken together, our results and those of Berent et al. (2021) show that repetition-based regularities are detected by infants in both linguistic and non-linguistic stimuli, but they engage different processing mechanisms. The dynamically increasing canonical hemodynamic response observed for linguistic stimuli suggests that the infant brain extracts the underlying structural regularity, i.e. reduplication, for these stimuli, but no such rule extraction takes place for tones or other non-linguistic stimuli, as indicated by the strong habituation, i.e. increasing deactivation we observed.

A large body of evidence suggests that the human brain is equipped with sophisticated abilities to efficiently process and organize perceptual input already from birth. Our study contributes to the exploration of these abilities. In an fNIRS study, we asked whether newborns could extract a repetition-based regularity from non-linguistic auditory stimuli, i.e., piano tones, like they do from speech. Our results show that newborns' ability to extract a repetition rule is not specific to language as we found differential brain responses to repetition- versus diversity-based sequences when those were carried by tones. However, the brain signatures of this response are different from those observed for structurally identical speech stimuli. We thus conclude that the detection of repetition is present for both linguistic and non-linguistic stimuli, but the underlying neural mechanisms differ.

As evidenced in this study, there are, in typically-developing newborns, specific brain mechanisms underlying the perception of speech and this may help them learn their native language. But what happens when speech perception is disrupted? How does this impact speech recognition and processing, which is fundamental for language learning? To better understand this, in a final experiment, we tested whether infants with hearing loss can discriminate their native language from a rhythmically different one, a task that typically developing newborns are able to perform (Ramus et al. 2000). To this effect, we compared infants with hearing loss and normal hearing control infants' brain responses to their native language, Italian, as well as to a rhythmically different unfamiliar language, English.

5 Speech recognition and processing in infants with hearing loss (Experiment 4)

5.1 Introduction

The journey to acquire our native language(s) starts even before our birth. During pregnancy, the auditory system of the fetus is operational from the 24th-28th week of gestation (Eggermont & Moore 2012). From then on, the fetus starts to experience speech. The prenatal speech signal is different from speech heard normally, because it is low-pass filtered by the womb tissues, which mostly only preserve the prosody, i.e. the melody and rhythm, of speech (Querleu et al., 1988; Gerhardt et al. 1992; Lecanuet & Granier-Deferre 1993; Griffiths et al., 1994) and filter out high frequencies necessary to identify individual sounds and thus words. At birth, newborns' listening preferences are shaped by this prenatal experience. Newborns prefer to listen to their mother's voice over an unfamiliar female voice (DeCasper & Fifer, 1980; Moon, 2017), to listen to a story that the mother read aloud during pregnancy compared to a novel story (DeCasper & Spence, 1986), and to their native language over a rhythmically different unfamiliar language (Mehler et al, 1988; Moon et al., 1993). This latter finding shows that newborns are already able to discriminate their native language (i.e., the language they heard in utero) from an unfamiliar language, based on rhythm, which is the means by which they experienced their native language prenatally.

More recently, brain imaging techniques have revealed how this prosodic prenatal experience already shapes how the newborn brain responds to language. At birth, the newborn brain is sensitive to the distinctive prosodic features of the language heard in utero (Abboub et al., 2016) and the native language already triggers a left-lateralized pattern of brain activation similar to what is found in adults (Peña et al., 2003; Kotilahti et al., 2009; Sato et al., 2011). One recent hypothesis (Gervain, 2015; Nallet & Gervain, 2021) proposes that the prosodic prenatal experience with the native language is fundamental for the discovery and later learning of language, as it shapes the auditory cortex to the rhythm of the native language, laying the foundation for further language acquisition. What happens then with speech perception abilities and language development when this prenatal experience is disrupted?

For instance, preterm babies born before the third trimester (during which the prosodic prenatal experience takes place) miss the prenatal experience with low-pass filtered speech and directly encounter the full-band speech signal. Many studies have highlighted the fact that children born preterm are at higher risk for developing language impairment (Barre et al., 2011), displaying delayed grammatical and vocabulary skills (Arpino et al., 2010; Guarini et al., 2016; Taylor et al., 2013; Vohr et al., 2014) as well as difficulties in learning to read and write (Guarini et al., 2009; Wolke et al., 2008). They also show disrupted speech perception abilities, affecting word stress discrimination (Herold et al., 2008) and even the neural processing of the native language seems to be atypical in the first months of life (Bartha-Doering et al., 2019). Bartha-Doering and colleagues (2019) compared preterm infants' brain responses to the native language at term-equivalent age to that of full-term neonates. While full-term neonates displayed the left-lateralized brain activation in response to the native language typical of adults, preterm infants at term equivalent age did not. These results suggest that missing the prosodic prenatal experience and prematurely encountering the full-band speech signal may affect how the infant brain processes the native language.

Another population who has atypical prenatal experience is infants with hearing loss. If the hearing loss is already present prenatally, this disrupts prenatal (and also postnatal) experience with speech. It is thus interesting to explore how infants with hearing loss perceive speech.

Hearing loss has different degrees or severity. Severity is assessed at different sound frequencies (500, 1000 and 2000 Hz, commonly) by measuring the intensity level in decibels at which a sound must be played for the patient to hear it. The threshold for normal hearing is equal to or below 20dB. Mild deafness corresponds to thresholds between 20 and 40 dB, moderate deafness between 40 and 70 dB, and severe deafness between 70 and 90 dB. A threshold above 90 dB corresponds to profound deafness. According to the World Health Organization, one and a half billion people in the world have some degree of auditory deficiency and more than 5% of the world's population is in need of readaptation service due to a disabling hearing loss, including 34 million children. Newborn hearing loss has various possible origins, both prenatal, like genetic factors or intrauterine infections, and perinatal, such as birth asphyxia, hyperbilirubinemia or low birth weight

(Biswas et al., 2012). Overall, there are more cases of hearing loss due to congenital origin than acquired deafness (i.e., postnatal hearing loss), highlighting the importance of neonatal hearing screening (Fortnum & Davis, 1997). A large survey published recently (Neumann et al., 2020) has investigated the early screening programs implemented in the world. One hundred and fifty-eight countries participated in the survey, from which only 41 had an early screening program testing at least 85% of the newborns in the country, primarily European and other developed countries, whereas Africa is the continent where early screening is the least common. The mean age at which infants are diagnosed with hearing loss is highly correlated with the presence or absence of screening at birth. If screened at birth, infants typically receive a diagnosis at 4,6 months, while without screening at birth, the diagnosis comes much later, at around 34,9 months, well into the age at which children should already produce language. This is crucial, as the importance of an early diagnosis for reducing language deprivation and optimizing good language outcomes has been highlighted in many studies (Yoshinaga-Itano et al., 1998; Kennedy et al., 2006; Pimperton et al., 2016). The survey (Neumann et al., 2020) also investigated which methods were used for neonatal hearing screening. The two main methods were otoacoustic emissions (OAE) used in 57% of the countries, and automated auditory brainstem responses (AABR) used by 11% of the countries, while 30% use a combination of these two methods.

Hearing loss affects auditory perception, and thus the perception of spoken language. Many studies have highlighted poor perception of prosodic cues by children with hearing loss, such as lexical and syllabic stress (Holt et al., 2016; Segal et al., 2016; Kalathottukaren et al., 2017; Konadath et al., 2021) or emotional prosody (Most & Michaelis, 2012). Moreover, prosodic perception has been shown to be a strong predictor of later language development in children with hearing loss (Grantham et al., 2022). However, very few studies to date have investigated the neural underpinnings of these abilities in young children with hearing loss.

In the present study, we investigated how hearing loss impacts the neural processing of speech within the first months of life. Specifically, we explored an early speech perception ability typically hearing newborns show: the discrimination of languages based on rhythm. For this purpose, we used a paradigm that has been commonly used with typically hearing newborns (Peña et al., 2003; Sato et al., 2011; May et al., 2015; 2017) and infants (Zhang et al., 2022) in order to study the brain responses to the native language and to an unfamiliar, rhythmically different language using fNIRS. In the first such study, Peña and colleagues (2003) presented Italian-exposed newborns with sentences in their native language played forward compared to the same sentences played backward and to silence while they recorded newborns' brain activity using fNIRS. Backward speech is used as a non-linguistic control stimulus, as it matches forward speech in all non-time-related characteristics (e.g. intensity, voice quality etc.). They found a significantly greater brain activation in response to the Forward condition, specifically localized in the left temporal area, suggesting the early presence of a left-hemispheric specialization for language processing in newborns.

To elucidate whether this pattern of activation was specific to the prenatally heard language or applies to all languages, Sato and colleagues (2011) used the same design but added an unfamiliar language condition. In this study, Japanese newborns were exposed to their native language, played both forward and backward, and to an unfamiliar language, English, played forward and backward as well. The Japanese Forward condition significantly activated a left temporo-parietal region, compared to Japanese backward and English forward. By contrast, no significant differences were found between the brain responses to forward and backward English, which both activated the temporal areas bilaterally. These results thus showed that the prenatally heard language already shaped the brain and triggered a left-hemispheric specialization in the newborn brain specifically for the processing of the native language.

The same design was then used by May and colleagues (2011) with English-exposed neonates. Participants listened to low-pass filtered stimuli (mimicking what newborns heard in utero) of their native language, forward and backward, and Tagalog as the unfamiliar language, backward and forward. Surprisingly, this study did not find left-dominant, but rather a bilateral activation, in response to the native language, although this activation was still greater overall for the native language English than for the unfamiliar language, Tagalog. The authors discussed these conflicting results and proposed two possible explanations, (i) this may be due to subtle differences in the probe placement between Peña et al. (2003) and their study, (ii) the use of low-pass filtered stimuli may have influenced processing, as Peña and colleagues used unfiltered speech stimuli. In a subsequent study with a similar design, May and colleagues (2017) tested the brain responses of English-learning newborns to English and Spanish. They found a large bilateral activation in the temporal area in response to the native language, and a similarly bilateral, but smaller activation to the unfamiliar one. As in Sato et al. (2011), they found a significant difference between forward and backward speech only for the native language. The discrimination of forward vs backward speech only being present for the native language was interpreted as a result of prenatal experience with the native language.

Using a slightly different design, Vannasing and colleagues (2016) conducted a NIRS study comparing the responses to forward and backward speech in the native and a non-native language in two separate groups of newborns. One group of French newborns was exposed to forward languages (French as the native language and Arabic as the non-native language), another group to the same languages but played backwards. The authors found a temporal, left-lateralized activation in response to the forward native language and a right-lateralized temporal response to the forward non-native language. Played backwards, both the native and the non-native language activated the bilateral temporal regions.

In summary, typically hearing newborns show a left-lateralized or bilateral temporal activation to the language they heard in utero and the discrimination between forward and backward speech depends on the familiarity with the language being heard.

To our knowledge, only one study used this paradigm with older infants (Zhang et al., 2022). As in Peña et al. (2003), Zhang and colleagues recorded the brain responses to the native language, forward and backward, without the addition of an unfamiliar language condition. In this study, two groups of Dutch infants were recruited. A group of 24 5-months-old infants and a group of 16 10-months-old infants. At 5 months, Dutch infants displayed a bilateral brain activation for both forward and backward speech, thus not displaying the left-lateralization found in Peña et al. (2003) and Sato et al. (2011) nor a difference between forward and backward speech (Peña et al., 2003; Sato et al., 2001; May et al., 2011; 2017). Also, responses were generally quite flat, i.e. small in amplitude. However, 10-month-olds did show a greater response to forward than backward speech in the left hemisphere, but these responses were inverted, i.e. hemodynamic responses whereby oxyHb concentrations decrease, while deoxyHb concentrations increase. Importantly, the newborn studies found a canonical positive hemodynamic response. Zhang and colleagues' (2022) study thus suggests that the neural processing of the native language may change within the first year of life. In particular, the observed pattern, i.e. canonical responses at birth, flat responses at 5 months and inverted responses at 10 months, may correspond to the increasingly easy and automatic processing of the native language as infants' experience and knowledge grows. More studies are needed to better understand this developmental trajectory.

Based on these findings with normal-hearing newborns and infants, we investigated the impact of hearing loss on the neural processing of speech in Italian infants with hearing loss (HL group) using a similar paradigm. We exposed infant participants to their native language (Italian) and to an unfamiliar, rhythmically different language (English), both in forward and backward speech while their brain activity was recorded using NIRS. A control group of normal-hearing infants (NH group) was also recruited. The control group was not perfectly matched in age or gender. Rather we recruited all infants between 0 and 12 months who were referred to the Audiology Ward of the University Hospital of Padua for hearing screening (on the basis of an initial positive hearing screening at birth or because of being at risk for hearing loss, e.g. due to genetic factors, preterm birth etc.). Some of these infants were diagnosed with a hearing loss, while others turned out to be normally hearing. We thus constituted two groups as a function of the outcome of the hearing screening at the Audiology Ward on the day of testing. Infants in the HL group had different degrees of hearing loss.

The exploratory nature of this study does not allow us to define precise hypotheses for the hearing loss group on the basis of existing results. However, we can still make some general predictions. First, we may hypothesize that the atypical auditory experience with speech in the HL group impacts how the brain processes languages. Thus, it is possible that we don't find significantly different responses to the native and the non-native language in the HL group, neither would we find a left-lateralized response to the native language. Also, we may expect to find a significant difference in the response to the Italian Forward condition between the HL group and the NH control group. Second, we can hypothesize that the degree of deafness will have an effect on the brain responses, as mild and moderate hearing loss may allow enough auditory experience to not impede language discrimination, while severe and profound hearing loss may disrupt it more strongly.

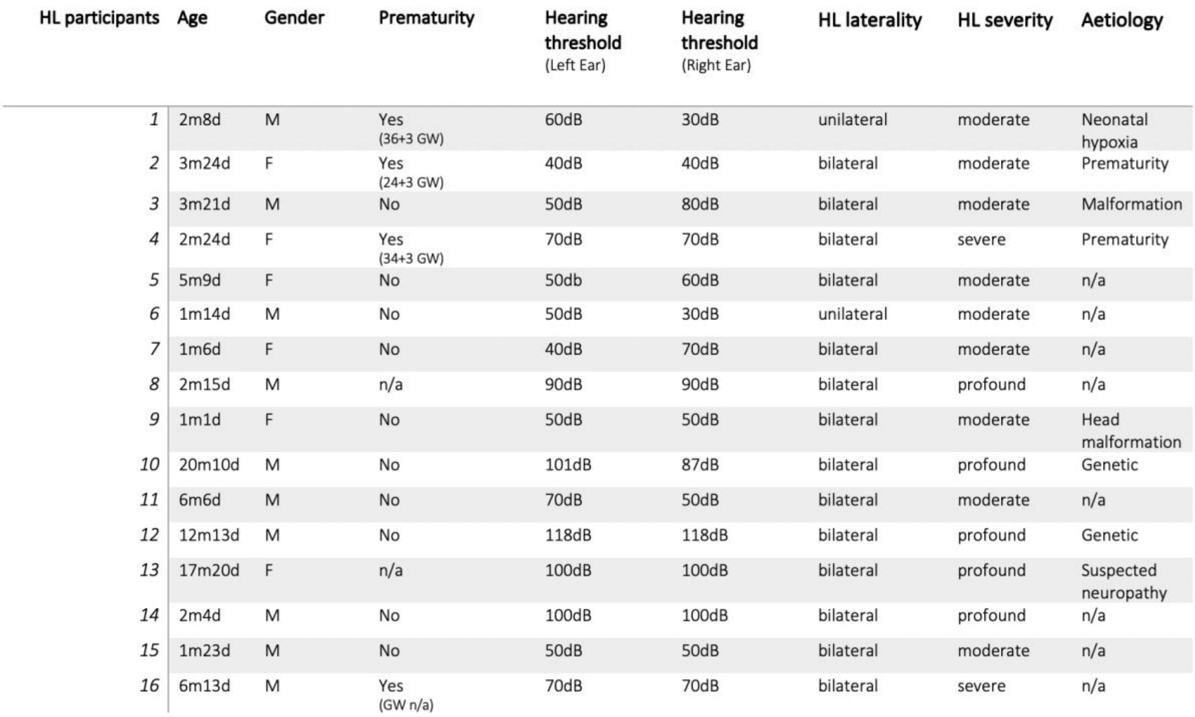
By contrast, we can formulate more precise hypotheses based on previous findings for the NH group (Peña et al., 2003; Sato et al., 2011; May et al., 2015; 2017; Vannasing et al., 2016; Zhang et al., 2022). Specifically, we hypothesize that the NH group will display a left-lateralized response to the native language played forward (Peña et al., 2003; Sato et al., 2011; Vannasing et al., 2016), but with a possible effect of age (Zhang et al., 2022).

5.2 Participants

Two groups of participants were recruited for this project: a group of infants who have been diagnosed with hearing loss and a group of normal hearing control babies. The infants were recruited at the Otolaryngology/Audiology Service of the University Hospital of Padua, Italy. All the participants had an appointment at the hospital to undergo medical hearing tests. Once all hearing tests had been completed, parents were invited to take part in the study with their infant. Depending on the final diagnosis, participants were assigned to the HL or the NH group. All parents of the infants tested gave written informed consent before the study, which was approved by the Ethics Committee for Clinical Trials (Comitato Etico per la Sperimentazione Clinica, CESC) of the University Hospital of Padua (protocol nr. AOP 2444). [The study was conducted in collaboration with Gaia Lucarini, University of Padua, who collected most of the infant data.]

5.2.1 Hearing loss group

Sixteen Italian infants with hearing loss, who had been exposed at least 50% of the time to Italian pre- and postnatally, were included in the final analysis (6 females; mean age: 173 days [approx. 5 months, 7 days]; age range: 32-618 days [approx. 1-20 months], mean birth weight: 2850g, birth weight range: 660-3640g). Twenty-eight additional infants were tested but excluded from the final analysis due to poor data quality (5), fussiness or crying (13) or less than 50% exposure to Italian (10). The infants in the final group had different degrees of deafness, originating from different aetiologies. Detailed information is presented in Table 5.1.

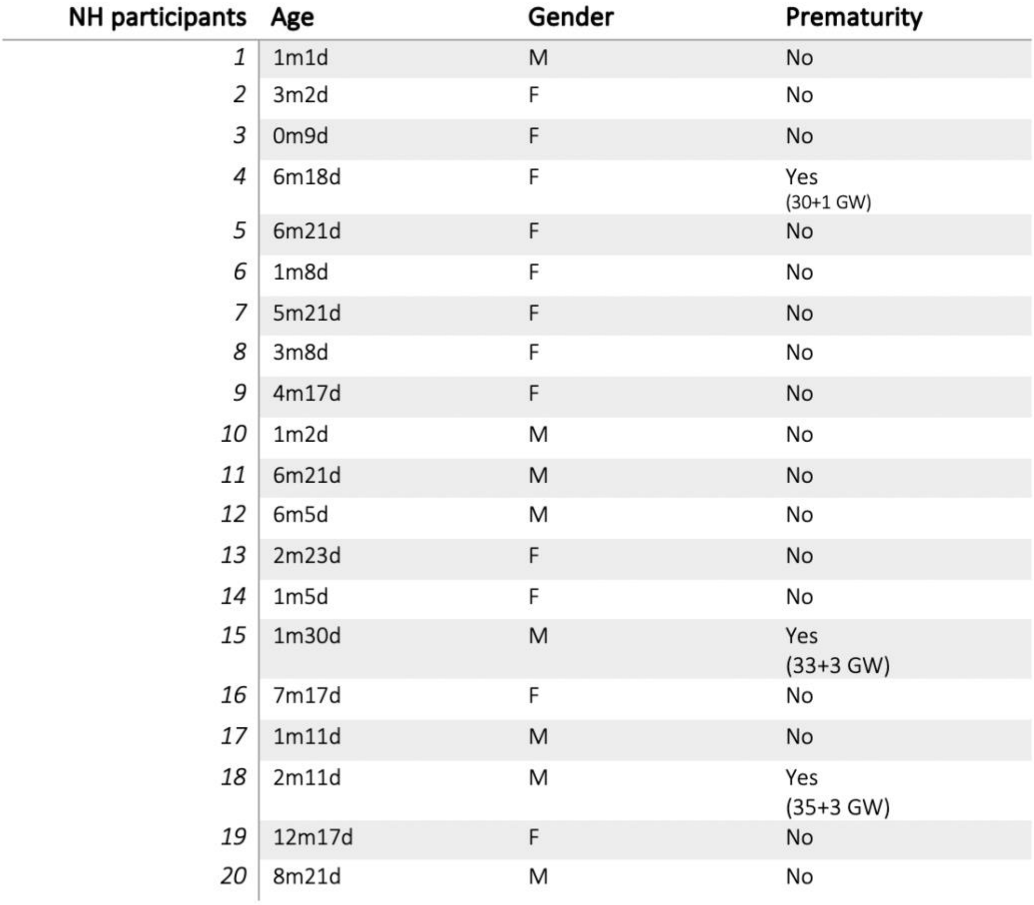


HL participants	Age	Gender	Prematurity	Hearing threshold (Left Ear)	Hearing threshold (Right Ear)	HL laterality	HL severity	Aetiology
1	2m8d	M	Yes (36+3 GW)	60dB	30dB	unilateral	moderate	Neonatal hypoxia
2	3m24d	F	Yes (24+3 GW)	40dB	40dB	bilateral	moderate	Prematurity
3	3m21d	M	No	50dB	80dB	bilateral	moderate	Malformation
4	2m24d	F	Yes (34+3 GW)	70dB	70dB	bilateral	severe	Prematurity
5	5m9d	F	No	50db	60dB	bilateral	moderate	n/a
6	1m14d	M	No	50dB	30dB	unilateral	moderate	n/a
7	1m6d	F	No	40dB	70dB	bilateral	moderate	n/a
8	2m15d	M	n/a	90dB	90dB	bilateral	profound	n/a
9	1m1d	F	No	50dB	50dB	bilateral	moderate	Head malformation
10	20m10d	M	No	101dB	87dB	bilateral	profound	Genetic
11	6m6d	M	No	70dB	50dB	bilateral	moderate	n/a
12	12m13d	M	No	118dB	118dB	bilateral	profound	Genetic
13	17m20d	F	n/a	100dB	100dB	bilateral	profound	Suspected neuropathy
14	2m4d	M	No	100dB	100dB	bilateral	profound	n/a
15	1m23d	M	No	50dB	50dB	bilateral	moderate	n/a
16	6m13d	M	Yes (GW n/a)	70dB	70dB	bilateral	severe	n/a

Table 5.1 : Demographic information of the Hearing Loss (HL) group. Age is expressed in months and days. Hearing thresholds are expressed in decibels. Abbreviations: m, months; d, days; dB, decibels; GW: gestational weeks.

5.2.2 Normal-hearing control group

Twenty Italian infants with normal hearing, who had been exposed at least 50% of the time to Italian pre- and postnatally, were included in the final analysis (13 females; mean age: 129 days [approx. 4 months]; age range: 9-382 days [approx. 0-12 months], mean birth weight: 3165g, birth weight range: 1600-3980g). Fifty-two additional infants were tested but excluded from analysis due to poor data quality (18), fussiness or crying (25) or less than 50% exposure to Italian (9). Detailed participant information is presented in Table 5.2.



NH participants	Age	Gender	Prematurity
1	1m1d	M	No
2	3m2d	F	No
3	0m9d	F	No
4	6m18d	F	Yes (30+1 GW)
5	6m21d	F	No
6	1m8d	F	No
7	5m21d	F	No
8	3m8d	F	No
9	4m17d	F	No
10	1m2d	M	No
11	6m21d	M	No
12	6m5d	M	No
13	2m23d	F	No
14	1m5d	F	No
15	1m30d	M	Yes (33+3 GW)
16	7m17d	F	No
17	1m11d	M	No
18	2m11d	M	Yes (35+3 GW)
19	12m17d	F	No
20	8m21d	M	No

Table 5.2 : Demographic information of the Normal Hearing (NH) group.

5.3 Stimuli

The stimuli consisted of (i) 16 Italian spoken sentences, (ii) the same Italian sentences played backward, (iii) 16 English spoken sentences, (iii) the same English sentences played backward. Some sentences were taken from the French subcorpora (parent-infant interaction in play situations) of the CHILDES (Child Language Data Exchange System) database, a large free database of children/parents' speech transcriptions (MacWhinney, 2000). Other sentences were taken from nursery rhymes or tales for children (The exhaustive lists of the English and Italian sentences are presented in Appendix A, in Table A.1 and Table A.2, respectively). The sentences were produced in an infant-directed manner by two Italian and two American-English female native speakers, each speaker contributed equally to the stimuli by producing eight sentences each. The sentences varied between affirmative, interrogative and exclamative. Sentences in the two language conditions were matched in duration; all the sentences contained 12 syllables. The sentences in the two languages did not differ significantly in duration, pitch or intensity. The backward sentences were created using the option “Reverse” in the Praat software (Boersma & Meenink, 2007) (Figure 5.2).

Figure 5.2 : The upper panel shows the waveform of the sentence “Lili always has breakfast before going to work” played forward. The lower panel shows the waveform of the same sentence played backward.

The experiment consisted of a total of 16 blocks, 4 blocks of each of the 4 conditions (Italian forward, Italian backward, English forward, English backward). Every block comprised 4 sentences. No sentence was repeated during the study. The order of the blocks was intermixed and counterbalanced across participants. No more than two blocks of the same condition occurred consecutively. Blocks were separated by silent intervals jittered between 15 and 18 seconds. Within blocks, sentences were also separated by silences jittered between 400 and 600 ms, yielding blocks of approximately 11 s. The whole experiment lasted about 8 minutes (Figure 5.3).

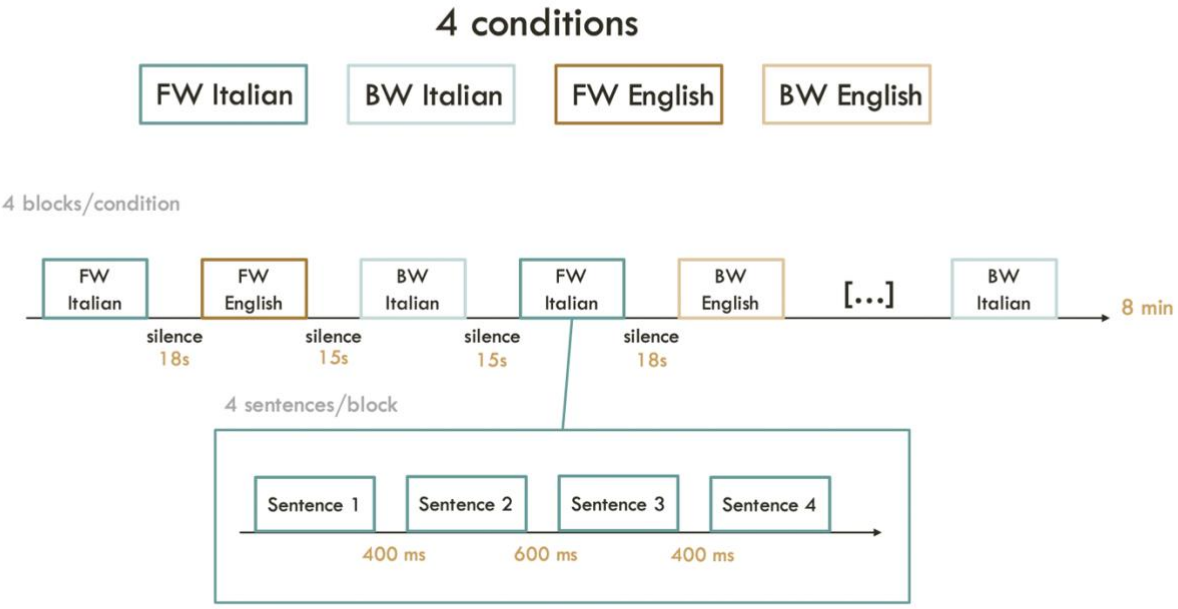


Figure 5.3 : Design of the study. The upper boxcar shows the organization and timescale of the entire experiment; the lower boxcar illustrates the organization of a single block.

To ensure sufficient data even when the infant participant doesn't complete the whole experiment, e.g. due to fussiness or crying, we organized the stimulus presentation such that two blocks of each condition were presented during the first half of the study, while the other two blocks of each condition were presented during the second half of the study (Figure 5.4). In this way, a participant who completed only half of the study anyway heard at least two blocks of each condition.

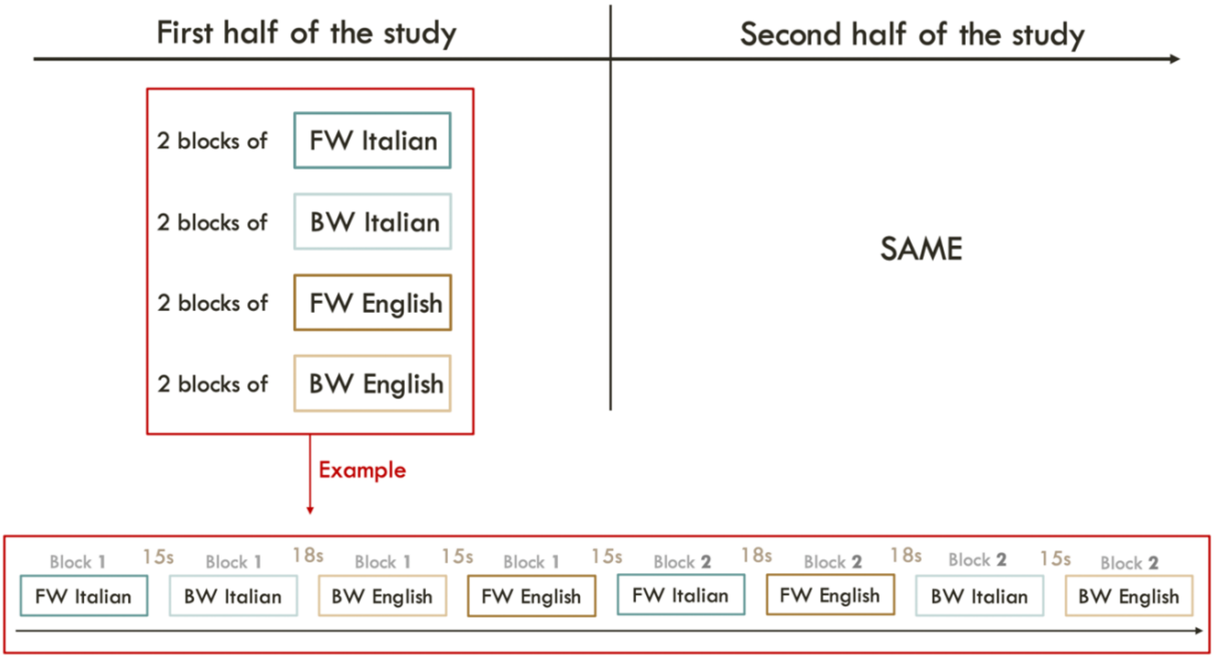


Figure 5.4 : Design of the whole experiment organized in two halves with two blocks of each condition presented in a randomized order during the first half the study, and the same organization for the second half.

5.4 Procedure

Infants were tested in a quiet room inside the Otolaryngology/Audiology Service of the University Hospital of Padua (Italy) or in a quiet, dimmed testing room at the Department of Developmental Psychology and Socialization, University of Padua, in the presence of at least one of their parents. During the study, infants were lying in a bassinette/pram (for the youngest ones) or were sitting on a parent's lap (Figure 5.5, panel A) while their brain responses were measured with a NIRx NIRSport2 8–8 machine (source-detector separation: 3 cm; pulsated LED lights at two wavelengths of 760 nm and 850 nm; sampling rate: 20.345 Hz). The optical probes were inserted into a stretchy cap (EasyCap, Brain-Products GmbH, Germany) placed on the infants' head using surface landmarks (nasion, and the preauricular points), covering the language areas in the bilateral temporal, frontal, and parietal cortices (10 channels/hemisphere, see Figure 5.5 Panel B). The size of the cap was chosen based on each infant's head circumference.

To localize the 20 channels, we projected their position relative to surface landmarks onto the cortical surface of 3D newborn anatomical head scans following Lloyd-Fox et al. (2014) and Shi et al. (2011), as in previous studies from our laboratory (Abboub et al., 2016; Martínez-Álvarez et al., 2022). Channels 1, 2, 4, 5 in the LH and 13, 14, 15, 16 in the RH were on average positioned over the frontal area, channels 3, 6 in the LH and 17, 19 in the RH over the temporal area, channels 9 in the LH and 21 in the RH were positioned in the frontier between the temporal and parietal areas, and channels 7, 10, 12 in the LH and 18, 20, 23 in the RH were on average positioned over the parietal area (see Figure 5.5, panel B).

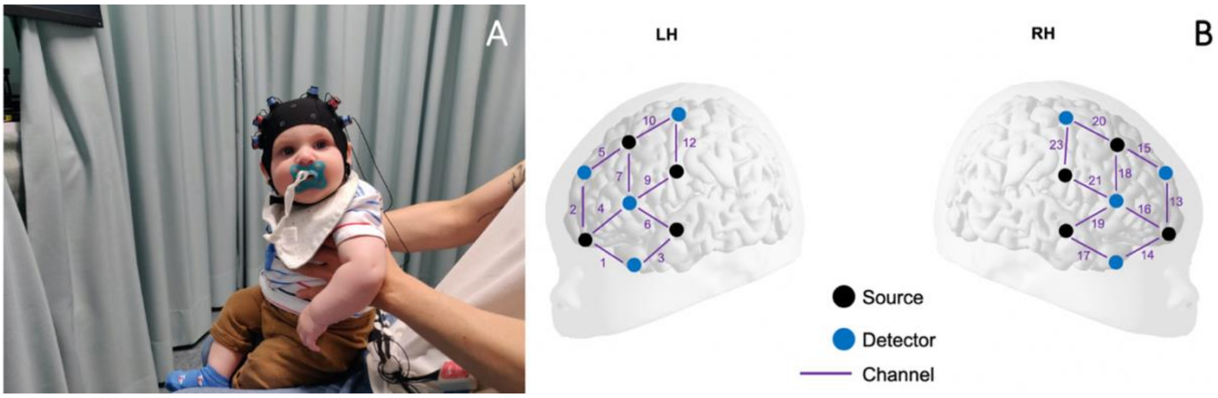


Figure 5.5 : The setup of the study and the position of the NIRS headgear. (A) A study participant with the headgear. (B) Configuration of light sources and detectors as well as NIRS measurement channels used in the study overlaid on a schematic newborn brain. Black circles indicate sources, while blue circles indicate detectors, and purple lines indicate channels.

Stimuli were presented using E-Prime, and delivered through two speakers placed at an angle of 30° on the two sides and approximately 1 m from the infant's head. The stimuli were presented at a conversational intensity of around 65dB. The computer running E-Prime sent time stamps to the NIRS machine.

In order to keep infants' attention during the study, those who were tested while seated in their parent's lap were facing a computer on which a cartoon suitable for young infants was played.

5.5 Data analysis and statistics

Very few studies exist on the neural correlates of basic speech perception abilities in infants with hearing loss, making our study exploratory. Thus, we conducted several statistical analyses.

First, we conducted in each group (HL and NH) separately a cluster-based permutation test based on a two-way ANOVA with Language (Italian/English) and Direction (Forward/Backward) as within-participant factors. One thousand permutations were conducted under the null hypothesis, for both oxyHb and deoxyHb. As post-hoc analyses, we conducted similar two-way ANOVAs with Language and Direction over the data in the time windows and channels of significance identified by the permutations. In order to understand how each condition activates the brain, we also ran cluster-based permutation analyses based on t-tests between each condition and the baseline in each group. Again, one thousand permutations were conducted under the null hypothesis, for both oxyHb and deoxyHb.

Secondly, as our samples vary considerably in age and the brain response to the native language has been found to change within the first year of life in normal hearing infants (Zhang et al., 2022), we analyzed the effect of the participants' age on their brain responses to the different conditions. To do so, in each group and for each condition, we extracted the oxygenated hemoglobin values from the bilateral temporal channels (3/6 and 17/19), typically activated by language in newborns and young infants (e.g. Experiment 1, Peña et al., 2003; Experiment 1, Gervain et al., 2008) in the entire time window of the brain responses (from the beginning of the stimulation to 20 seconds after the end of the stimulation) and we performed a linear regression with age as the covariate.

Thirdly, to better understand how the severity of deafness impacts brain responses in the HL group, we followed the same procedure as for the factor age, but with the hearing thresholds as the covariate. Analyses were carried out by taking into account the hearing threshold in the right ear only, the left ear only and the average threshold for both ears.

Fourthly, we examined the influence of hearing loss by directly comparing the responses of the two groups (HL/NH) to Italian Forward in the predefined bilateral temporal channels (channels 3 and 6 on the left hemisphere and channels 17 and 19 on the right hemisphere), as well as in the channels that emerged in the permutation tests.

5.6 Results

The grand average hemodynamic responses of the two groups are presented in Figure 5.6 (Panel A: Hearing Loss group; Panel B: Normal-Hearing group).

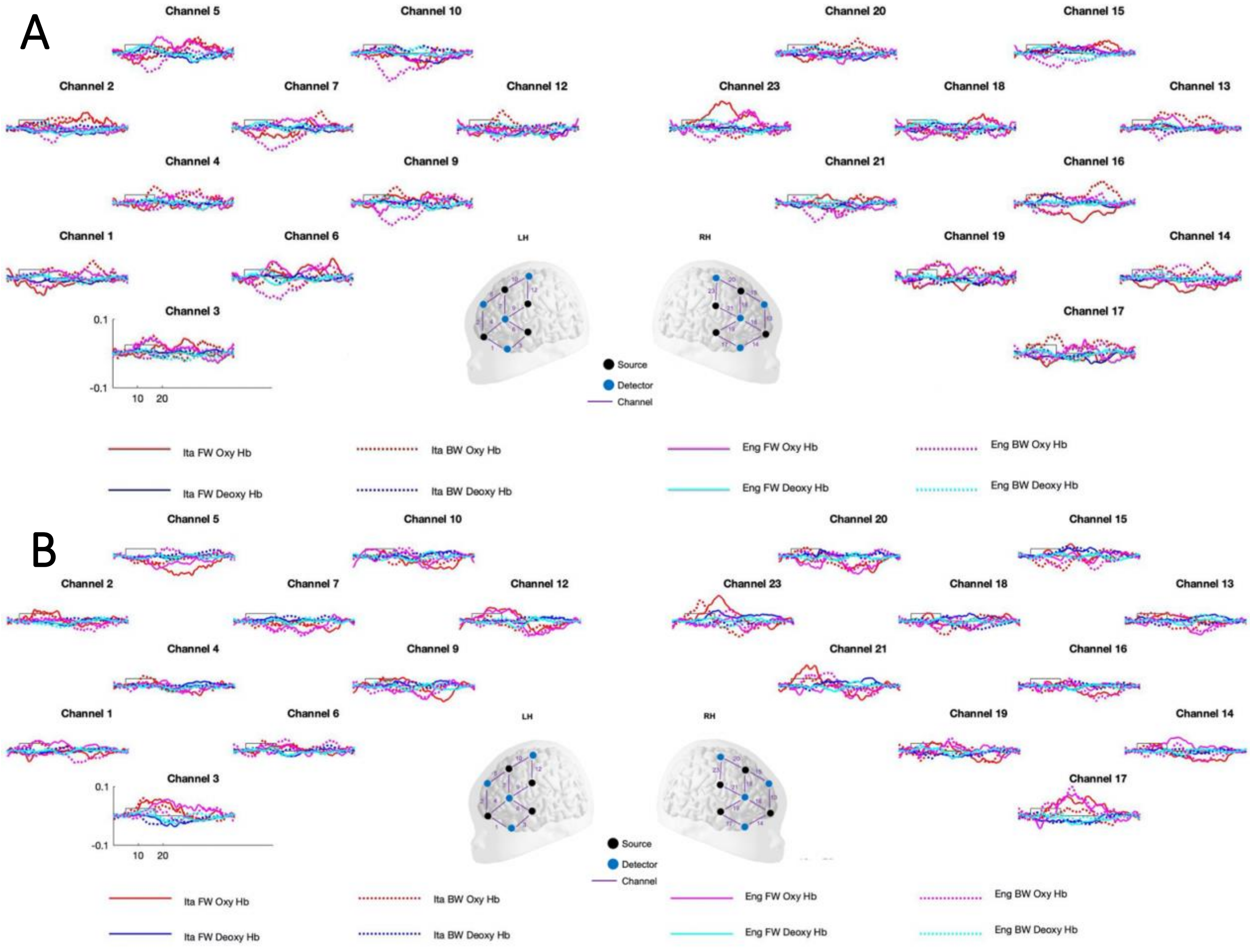


Figure 5.6 : The grand average hemodynamic response of the hearing loss (HL) group (A) and the normal hearing (NH) group (B). Channels are plotted following the probe placement displayed in Figure 5.5. The x-axis represents time in seconds; the y axis shows the concentration in mmol x mm. The rectangle along the x-axis indicates the time of stimulation. OxyHb and deoxyHb concentration changes in response to the Italian Forward condition are shown in red and blue, respectively (continuous lines). OxyHb and deoxyHb concentration changes in response to the Italian Backward condition are shown in red and blue, respectively (dashed lines). OxyHb and deoxyHb concentration changes in response to the English Forward condition are shown in magenta and cyan, respectively (continuous lines). OxyHb and deoxyHb concentration changes in response to the English Backward condition are shown in magenta and cyan, respectively (dashed lines).

5.6.1 Brain responses to speech in the Hearing Loss group

OxyHb.

The cluster-based permutations with a two-way ANOVA revealed a main effect of Language (Italian vs English) in the bilateral parietal region (channels 10 and 12 in the left hemisphere (p<.001) and channel 18 in the right hemisphere) (p<.001), a main effect of

Direction (Forward vs Backward) in a right frontal channel (channel 14) (p<.001) and a significant interaction Language X Direction in the right parietal region (channels 20 and 23) (p<.001). The mean responses in the significant clusters for the main effects are presented in Figure 5.7. The main effect of Language is due to the fact that both languages triggered negative (inverted) responses in the left parietal area with a more negative response to English than to Italian and a greater positive activation for Italian in the right parietal area (Figure 5.7 Panel A). The main effect of Direction was carried by a greater activation for backward than forward speech in the right frontal region, with backward speech giving rise to a positive hemodynamic response, forward speech a negative one (Figure 5.7 Panel B). To understand what carries the significant interaction Language X Direction in the right parietal region, we extracted the oxyHb values from this cluster in the significant time window and performed a two-way ANOVA with Language and Direction as within-subject factors. In this analysis, the interaction Language X Direction was not significant (p=.787), which is in conflict with what was found with the cluster-based permutation ANOVAs. But since the interaction was significant in the permutations, we report here the pairwise comparisons: t-tests between each conditions pair revealed no significant comparison. The two comparisons approaching the most significance (p=0.163 for both) are Italian Backward vs English Forward and Italian Backward vs English Backward. The corresponding mean responses are presented in Figure 5.8. The discrepancy between the results of the two analyses is explained by the fact that the permutations compute statistics over the entire dataset, i.e. all data points from all conditions and channels are permuted, whereas this ANOVA is specific to the much smaller dataset of the originally significant cluster and time window. The discrepancy nevertheless suggests that the effects found in the cluster-based permutation ANOVA are likely weak, possibly due to the small sample size, and must be interpreted with caution.

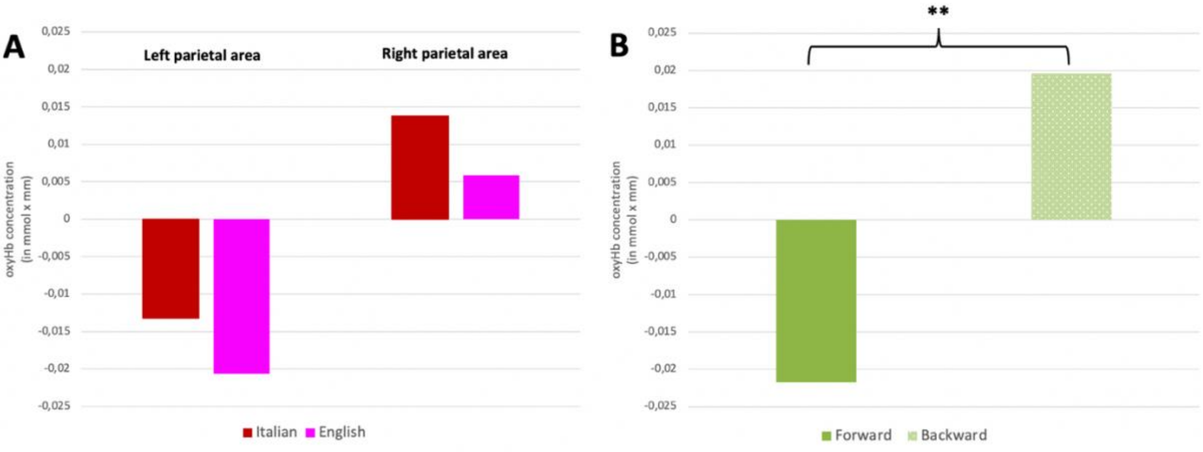


Figure 5.7 : (A) Mean oxyHb concentration in response to Italian and English conditions in the left and right parietal area. (B) Mean oxyHb concentration in response to Forward and Backward conditions in the right frontal area.

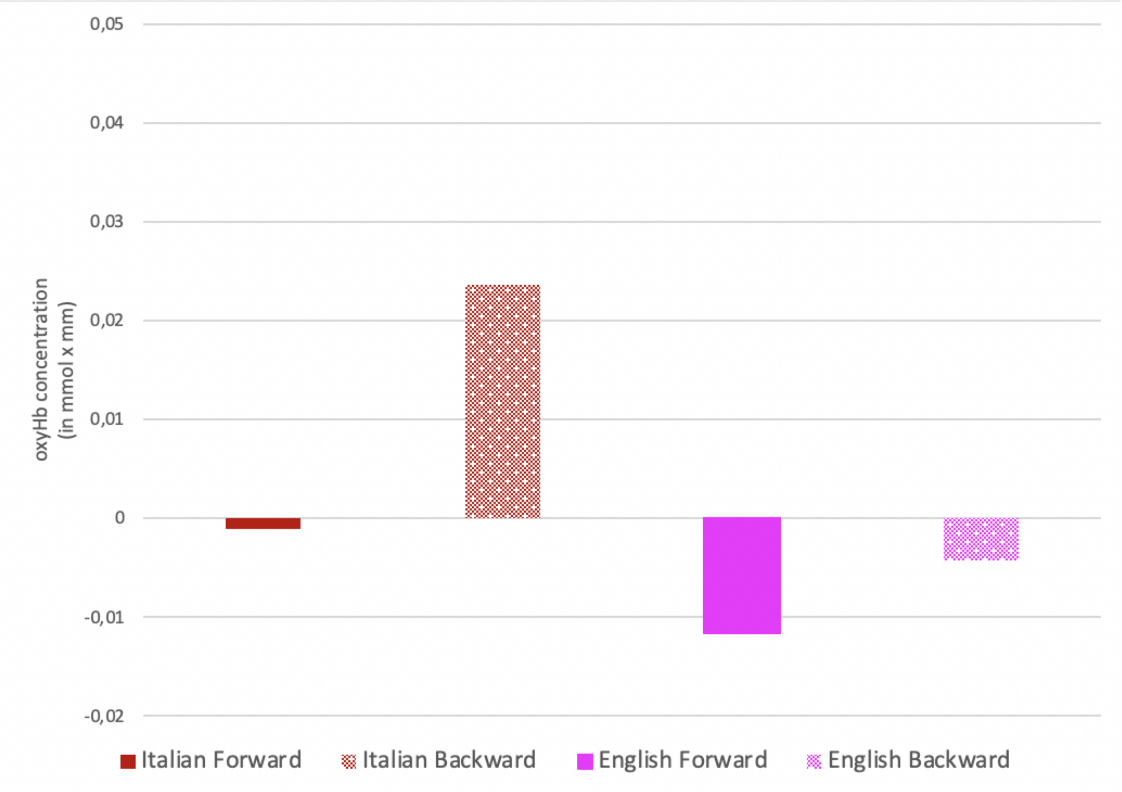


Figure 5.8: Mean responses to the four conditions in the right parietal area.

DeoxyHb.

The cluster-based permutation analysis with a two-way ANOVA on deoxygenated hemoglobin revealed a main effect of Language (Italian/English) in a left parietal channel

(channel 7) (p.<.001), a main effect of Direction (Forward/Backward) in a left parietal channel (channel 7) (p.<.001) and a right frontal cluster (channels 13, 14 and 16) (p.<.001), and a significant Language X Direction interaction in a large left frontal channel (channel 5) (p<.001) and in a right parietal channel (channel 20) (p<.001).

The statistical comparisons between each condition and the baseline revealed the following.

OxyHb

Italian Forward vs Baseline. The cluster-based permutation analysis highlighted a significant canonical hemodynamic response in a right parietal channel (channel 23) (p<.001) compared to baseline.

Italian Backward vs Baseline. Italian backward triggered significant canonical activation in the bilateral frontal areas (channels 1, 2 and 4 in the left hemisphere and channels 13 and 16 in the right hemisphere) (p<.001).

English Forward vs Baseline. No channel was significantly activated by English Forward compared to baseline.

English Backward vs Baseline. English backward triggered a significant inverted response bilaterally in the left temporal (channels 6 and 9) (p<.001) and left fronto-parietal areas (channels 5, 7 and 10) (p<.001) and the right parietal area (channels 21 and 23) (p<.001).

DeoxyHb

Italian Forward vs Baseline. The deoxy response was significantly different than the baseline in a left parietal channel (channel 10) (p<.001).

Italian Backward vs Baseline. No channel showed significant activation in this comparison.

English Forward vs Baseline. The analysis showed a significantly different response from the baseline in a left frontal region (channels 2 and 5) (p<.001) due to a canonical response and in the bilateral fronto-parietal area (channels 5 and 7 on the left hemisphere and channels 15, 16 and 18 on the right hemisphere) (p<.001) due to an inverted response.

English Backward vs Baseline. English backward triggered a significant inverted deoxy response compared to baseline in a left fronto-temporal area (channels 4 and 6) (p<.001) and in a right parietal region (channels 21 and 23) (p<.001).

5.6.2 Brain responses to speech in the Normal Hearing group

OxyHb.

The cluster-based permutation analysis with a two-way ANOVA revealed a main effect of Language (Italian/English) in a large left fronto-temporo-parietal cluster (channels 1, 2, 4, 6 and 7) (p.<.001) and a large right fronto-parietal cluster (channels 14, 16 and 17) (p.<.001), a main effect of Direction (Forward/Backward) in a left fronto-temporal cluster (channels 6 and 9) (p.<.001) and a large right fronto-temporo-parietal clusters (channels 15, 18, 20 and 21) (p.<.001), and a significant Language X Direction interaction in a left fronto-temporal cluster and in a right temporo-parietal channel (channel 21). The mean responses in the significant clusters for the main effects are presented in Figure 5.9. The main effect of Language was due to the fact that Italian triggered positive responses, while English triggered negative (inverted responses) in the left fronto-temporo-parietal area while in the right fronto-temporal area, both languages triggered positive responses with greater activation to English than Italian (Figure 5.9 Panel A). The main effect of Direction was due to forward speech triggering more activation than backward speech, with an inverted response in the left ROI and to a canonical response in the right ROI (Figure 5.9 Panel B). To understand what triggers the significant Language X Direction interaction in the left fronto-temporal region and the right temporo-parietal channel (21), we extracted the oxyHb values in the significant time windows and channels and performed a two-way ANOVA with Language and Direction. This analysis revealed that the interaction Language X Direction was marginally significant (p=0.083) in the left hemisphere, and significant in the right hemisphere (p=0.015).Post-hoc t-tests between each condition revealed no significant comparison but two comparisons approach significance, the difference between English Forward and English Backward (p=0.085) and between English Forward and Italian Forward (p=0.087). We may infer that with a bigger sample these comparisons would be significant. The corresponding mean responses are presented in Figure 5.10.

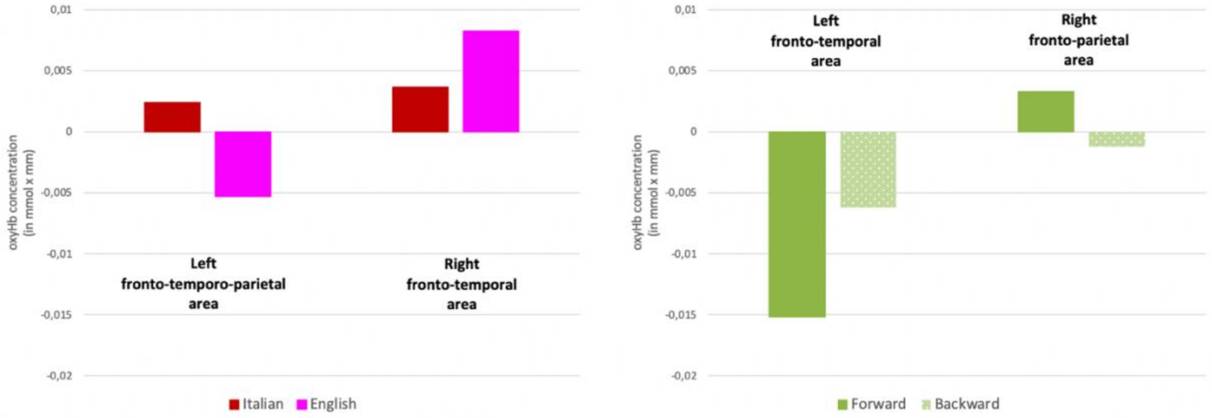


Figure 5.9 :(A) Mean oxyHb concentration in response to Italian and English conditions in the left fronto-temporo-parietal area and the right fronto-temporal area. (B) Mean oxyHb concentration in response to Forward and Backward conditions in the left fronto-temporal area and the right fronto-parietal area.

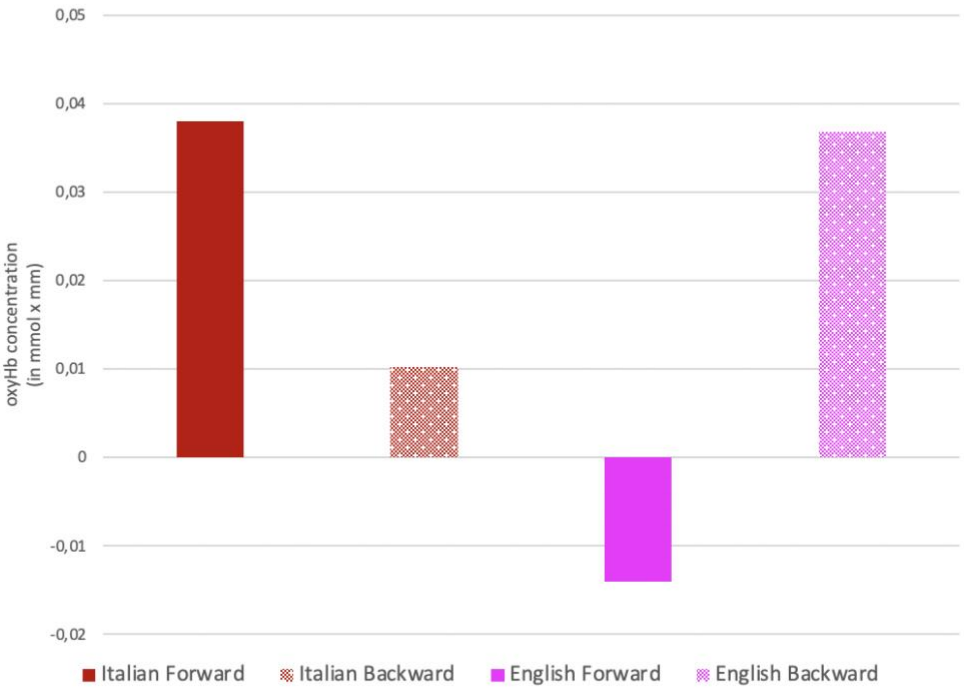


Figure 5.10 : The mean responses to the four conditions in the right temporo-parietal area.

DeoxyHb.

The cluster-based permutation analysis with a two-way ANOVA on deoxygenated hemoglobin revealed a main effect of Language (Italian/English) in a left frontal channel

(channel 4) (p.<.001) and a large right fronto-temporal cluster (channels 14, 17 and 19) (p.<.001), a main effect of Direction (Forward/Backward) in a left parietal channel (channel 7) (p.<.001) and a right frontal channel (channel 14) (p.<.001), and a significant Language X Direction interaction in a large left fronto-temporo-parietal cluster (channels 2, 4, 6 and 7) (p<.001) and in a right frontal cluster (channels 14 and 16) (p<.001).

The statistical comparisons between each condition and the baseline revealed the following.

OxyHb

Italian Forward vs Baseline. The cluster-based permutation analysis revealed that Italian Forward triggered a significant inverted response in a left fronto-parietal region (channels 5, 7, 9 and 10) (p<.001).

Italian Backward vs Baseline. No channel was significantly activated compared to baseline by the Italian Backward condition.

English Forward vs Baseline. The analysis yielded a significant canonical activation in the bilateral fronto-temporal area (channels 1 and 3 in the LH and channels 14 and 17 in the RH) (p<.001), as well as a left parietal activation (channels 10 and 12) (p<.001), which was also canonical.

English Backward vs Baseline. The analysis revealed a significant inverted response to English backward in the left parietal area (channel 12) (p<.001).

DeoxyHb

Italian Forward vs Baseline. The analysis revealed a significant response in the left frontal region (channel 13) (p<.001) and the left parietal region (channels 21 and 23) (p<.001).

Italian Backward vs Baseline. The analysis yielded a bilateral fronto-temporal region where responses differed significantly from the zero baseline (channels 5 and 10 in the LH and channel 17 in the RH) (p<.001).

English Forward vs Baseline. The cluster based permutations highlighted a right temporal region where the response differed significantly from the baseline (channels 17 and 19) (p<.001)

English Backward vs Baseline. A left frontal cluster showed a significantly different response from the zero baseline (channels 2 and 4) (p<.001).

5.6.3 Influence of the age on the brain responses

In the HL group, the analysis showed a significant effect of Age on the responses to Italian Backward in the left temporal area (p=0.043), explained by an increasing activation with age (Figure 5.11 Panel A). Age did not influence the brain responses to the other conditions, nor to the Italian Backward condition in the right temporal area.

In the NH group, the analysis revealed a significant effect of Age on the responses to Italian Forward in the left temporal region (p=0.041), explained by an increasing activation with age (Figure 5.11 Panel B). Age did not influence the brain responses to the other conditions, nor to the Italian Forward in the right temporal area.

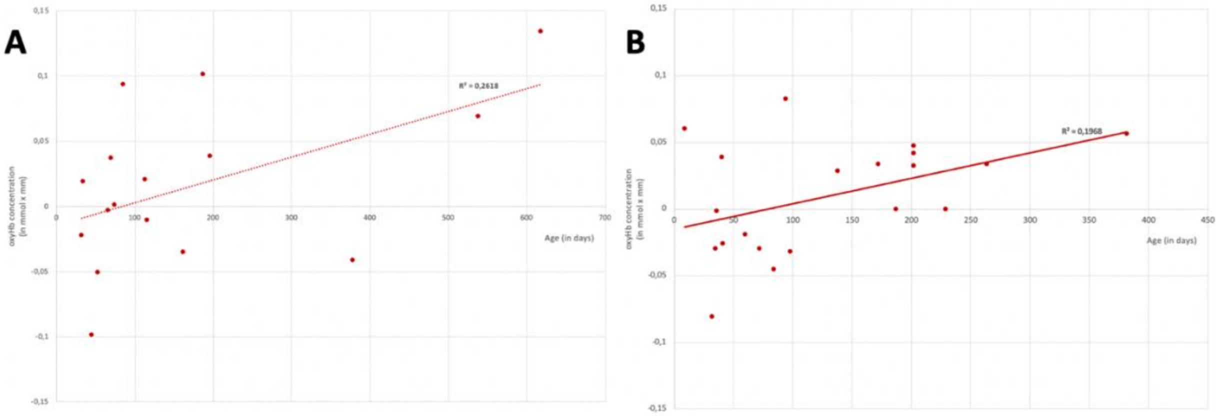


Figure 5.11 : (A) Brain responses to Italian Backward in left temporal area for the HL group as a function of age. (B) Brain responses to Italian Forward in the left temporal area for the NH group as a function of age.

5.6.4 Influence of the hearing thresholds on the brain responses

The analyses with the hearing thresholds of the right ear, the one sending afferents to the LH, where the language network is located in the majority of the population, and of the averaged thresholds of the two ears did not reveal any significant effect of the thresholds on the brain responses to the different conditions. The hearing thresholds of the left ear showed a marginally significant effect on the brain responses to English Backward in the right temporal area (p=0.066) (Figure 5.12).

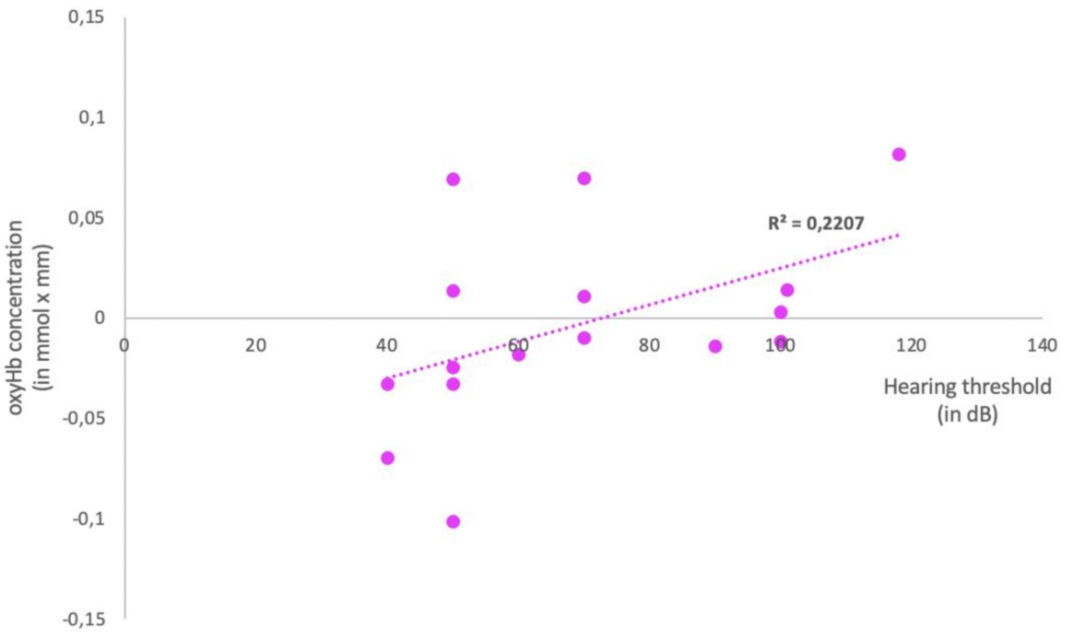


Figure 5.12 : Effect of the hearing threshold in the left ear on the brain responses to English Backward in the right temporal area.

5.6.5 Direct comparison between groups

To better examine the impact of hearing loss, we compared responses to the native language, Forward Italian between the two groups in the bilateral temporal areas (channels 3 and 6 in the LH and channels 17 and 19 in the RH). The analysis revealed no significant differences between the hearing loss and the control group.

We then looked at group differences within the clusters highlighted by the permutations. Compared to baseline, Italian Forward significantly activated a left fronto-parietal cluster (channels 5, 7, 9 and 10) in the NH group and a right parietal channel in the HL group (channel 23). We extracted the data in the two groups from these two clusters in the time windows of significance and performed independent samples t-tests. The analysis revealed a significant effect of the group on the response to Italian Forward in the left fronto-parietal area (p=0.033) (Figure 5.13 Panel A), but no difference in the right parietal area (Figure 5.13 Panel B). The significant difference between the two groups in the LH was due to a stronger inverted response in the NH group.

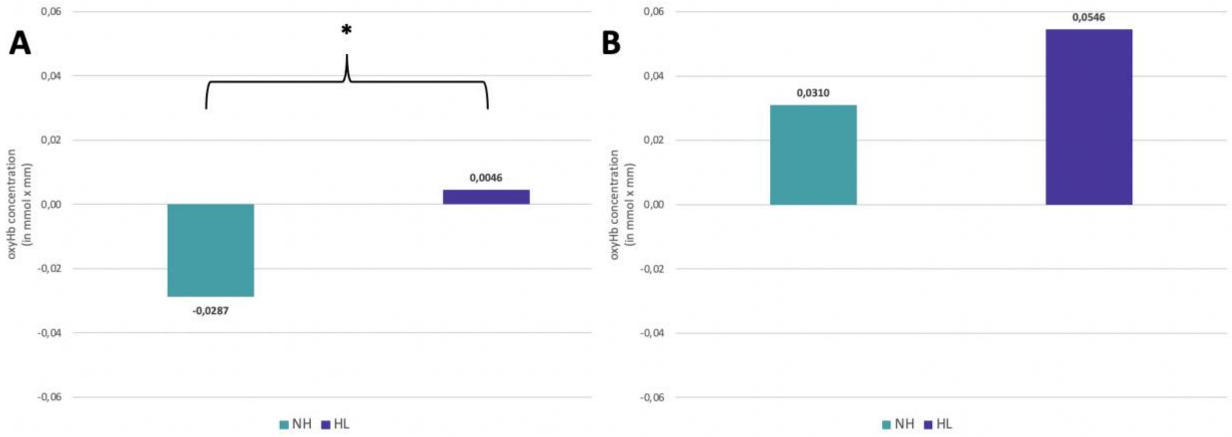


Figure 5.13: Brain responses to the native language in the NH and HL groups in the left fronto-parietal region (A) and in the right parietal region (B).

5.7 Discussion

In this study, we investigated how hearing loss may impact the early brain processing of language in infants. To do so, we implemented a paradigm commonly used in typically hearing newborns (Peña et al., 2003; Sato et al., 2011; May et al., 2015; 2017; Vannasing et al., 2016) and infants (Zhang et al., 2022) to investigate how their brains processes the native language and an unfamiliar language. In our study, we exposed Italian-learning infants with hearing loss (HL group) and control normal-hearing infants (NH group) to their native language, Italian and to an unfamiliar, rhythmically different language, English. Both languages were played forward in half of the blocks and backward in the other half. The brain responses of the infants were measured using NIRS in the frontal, temporal and parietal areas bilaterally.

Our study was exploratory as this paradigm had never been used in infants with hearing loss before and in only one study with normal-hearing infants older than newborns (Zhang et al., 2022). Studying clinical populations is challenging as it often involves heterogeneous groups and low sample sizes (Prince et al., 2021). This is no different for infants with hearing loss, who may have various degrees of hearing loss, from various aetiologies, accompanied or not with comorbidities. Indeed, the HL group in our study had a relatively low sample size (16 participants) and was heterogeneous both in the severity and laterality of the hearing loss and in age (between 1 and 20 months of age). Moreover, we obtained some conflicting results between different statistical analyses, suggesting that our effects are weak. In the HL group, some results were not congruent between oxyHb and deoxyHb concentration changes, that may reflect unclear, non-canonical hemodynamic response in some channels for some conditions. These parameters imply that our results must be taken with caution and that more focused studies are needed in the future to allow stronger conclusions. This being said, we obtained some noteworthy results.

First, in the NH group, our results confirm the hypothesis that typically hearing infants show a left-lateralized activation for the native language. Indeed, the two-way ANOVA and the comparison with baseline both highlighted a significant response to Forward Italian in a left fronto-temporo-parietal areas. While newborns data commonly show a positive hemodynamic response to the native language, we found a significant inverted response to Italian Forward. This is in line with what Zhang and colleagues (2022) found in 5- and 10-month-old infants, i.e. ages similar to our sample. Our second hypothesis concerning the NH group was that we may find an effect of the age on the brain responses to the native language as Zhang and colleagues found different patterns of results between the 5-month and the 10 month groups. In their NIRS study, the younger group did not show the left-lateralized activation for the native language played forward, but forward and backward speech triggered bilateral activation. However, the older group did show a greater, left-lateralized, response to forward Dutch compared to backward Dutch, but this response was inverted. As discussed in the introduction, this developmental pattern could be due to a linear decrease of the brain response to the native language, which is positive at birth, flat at around 5-months and inverted at the end of the first year. Another possibility would be a U-shape developmental trajectory of the brain activation to the native language during the infant's first year, considering the activation of the left hemisphere, regardless of the direction of the activation (positive vs negative hemodynamic response). At birth, newborns show a significant positive hemodynamic response which is linked to the prenatal experience with the native language that shapes speech perception at birth. A few months later, the infant grows in a more diverse environment, which could lead to a less specific response for the native language in particular. Linked to that, an interesting study (de la Cruz-Pavía & Gervain, 2023) showed that while newborns show greater brain activation for repetition-based structure (AAB; Gervain et al., 2008; 2012), at 6-months, infants display greater activation for diversity-based sequences (ABC). The authors linked this brain particular brain activation to word learning that starts at this age and require the processing of difference in the linguistic input. At the end of the first year, infants had a greater experience with the native language and the language acquisition intensifies, which may be reflected by a significant brain activation for the native language again. However, this latter explanation does not explain why the direction of the brain response (positive at birth and negative at the end of the first year) changes within the first year of life, while the first explanation does. Confirming our hypothesis, our results revealed a significant effect of Age on the responses to Italian Forward in the left temporal region. This is in line with Zhang and colleagues' findings. More studies measuring the brain responses to the native language in infants of various ages within the first year of life are needed to characterized more precisely this developmental trajectory.

Furthermore, the unfamiliar forward language triggered bilateral activation. The comparison with the baseline highlighted a bilateral fronto-temporal activation to English Forward and the two-way ANOVA revealed a main effect of Language with a greater activation for the native language in the left hemisphere but in the right fronto-parietal area for the unfamiliar language. This bilateral activation for the unfamiliar language has already been found in newborns (Sato et al., 2011; May et al., 2017). In one study (May et al., 2017), both the native and the non-native language triggered bilateral temporal activation. Our results are thus more in line with Sato and colleagues' results (2011), who found a significant activation in the right hemisphere for the non-native language only. This right-lateralization may be explained by a discrimination of the native and non-native language based on rhythm, as it has been shown that the right hemisphere is linked to speech rhythm processing (Geiser et al., 2008).

Finally, the NH group displayed overall greater activation for forward than backward speech, i.e. the classical forward-advantage found in most other studies for the native language (Peña et al., 2003; Sato et al., 2011), which may suggest the presence of a greater brain activation for linguistic stimuli compared to non-linguistic counterparts, irrespective of familiarity.

In summary, the results of our control group are in line with the literature, showing a left-dominance for the native language influenced by age (Zhang et al., 2022), as well as a greater activation for forward than backward speech (Peña et al., 2003; Sato et al., 2011, Zhang et al., 2022).

Results in the HL group showed a qualitatively different pattern, although we need to interpret our results with caution, as discussed earlier. For this group, we hypothesized that hearing loss may impact the neural processing of language, possibly leading to the absence of specific left-lateralized activation in response to the native language. Indeed, our results did not show an activation of the left hemisphere in response to Italian Forward, rather a right-lateralized activation. The comparison with the zero baseline revealed a significant activation of the right parietal region triggered by Italian Forward, while the two-way ANOVA highlighted a main effect of Language, in particular in the right hemisphere due to a larger response to Italian than English. This pattern of results suggests that reduced auditory experience with the native language impedes the emergence of the usual left dominance for the native language. The right-lateralized activation in response to the native language in this group may be linked to a coarse, imprecise perception of the native language based on rhythm only, even after birth, due to a reduced perception of smaller, more fine-grained linguistic units (phonemes, words, etc.).

Interestingly, for Direction, the HL group showed the opposite pattern than the NH group, with a greater overall activation for Backward than Forward speech. The two-way ANOVA revealed a significant effect of Direction in the right frontal area with forward speech triggering an inverted response while backward speech eliciting a canonical hemodynamic response. This effect is robust, as it has been confirmed by the post-hoc repeated measures ANOVA. Moreover, the comparison with the baseline highlighted one significant channel in response to Italian Forward and no significant activation in response to English Forward while the backward conditions elicited large bilateral activations. Italian Backward activated the bilateral frontal area, while English Backward triggered significant inverted response bilaterally, in the left temporal and left fronto-parietal region and in the right parietal region. This greater activation for backward speech in infants with hearing loss may be linked to their atypical auditory experience. Hearing loss renders auditory stimuli more complex to process and this may result in more complex, less familiar stimuli triggering greater brain activation than less complex, more familiar ones.

When we examined the effect of hearing thresholds, we found a marginally significant effect of the left ear’s hearing thresholds on the response to English Backward in the right temporal area. However, it seems that the effect is carried by the few profoundly deaf babies (hearing thresholds equal to or over 100 dB), who could arguably hear very little of the auditory stimuli. How is the possible then that these infants had large positive hemodynamic responses in the temporal area? We speculate that this result is linked to a potential cortical reshaping due to auditory deprivation, and thus may represent infants’ responses to the visual attention getters we used. Indeed, if the auditory cortex is not stimulated by auditory stimuli, it may be recruited for the processing of other modalities, like visual processing. This cross-modality cortical reshaping has been documented in profoundly deaf adults showing a temporal activation in response to visual stimulation (Dewey & Hartley, 2015), deaf children (Shi et al., 2015) and even deaf infants (Xia et al., 2017). Interestingly, Dewey and Hartley (2015) used NIRS to investigate cross-modal reshaping in deaf adults. They recorded the brain activity of the temporal lobes to auditory, visual and somatosensory stimuli in profoundly deaf adults and in typically hearing controls. They found that visual stimulation elicited significant activation in the right temporal area in deaf participants but not in controls. This study suggests that NIRS is a good tool to investigate cortical reshaping. Moreover, Xia and colleagues' results (2017) suggest that cortical reshaping may appear already during infancy. Taken together, these findings suggest that our NIRS data may already reveal a cortical reshaping of the auditory cortex in the profoundly deaf infants of our sample. Of course, more studies using a specific visual stimulation paradigm and a larger, homogeneous group of profoundly deaf infants are needed to confirm and better explore early cortical reshaping in deaf infants.

Finally, we directly compared brain responses to Italian Forward in the two groups, and found significant differences in the left fronto-parietal region. The native language triggered almost flat responses in the HL group, while the NH group showed a large inverted response. This significant difference suggests that hearing loss impacts how the brain processes the native language and impedes left-hemispheric specialization for the native language found in typically-hearing newborns (Peña et al., 2003, Sato et al., 2011; Vannasing et al., 2016), infants (Deheane-Lambertz et al., 2002) and adults (Molfese et al., 1975; Ojemann et al., 1989; Binder et al., 1995; Dehaene et al., 1997; Knecht et al., 2000; Olulade et al., 2020).

To sum up, we investigated the neural processing of the native language and an unfamiliar language in infants with hearing loss and normally-hearing controls. The control group exhibited the left-hemispheric specialization for the processing of the native language well documented in several previous studies with typically developing infants (Dehaene et al., 2002; Peña et al., 2003, Sato et al., 2011; Vannasing et al., 2016). By contrast, HL infants showed atypical language processing, with a weaker, right-lateralized brain activation for the native language and a significant sensitivity of the brain to complex stimuli (backward speech). These results uncover the strong and developmentally very early impact of hearing loss on spoken language processing and the organization of the neural language network.

6 General Discussion

In this research work, we have investigated the early neural processing of speech and other linguistic and non-linguistic auditory stimuli in newborns and young infants. The general aim of this thesis is to better understand how the newborn and infant brain processes speech and whether the processing of speech is specific compared to other auditory stimuli, linguistic or non-linguistic, communicative or not, produced by humans or not.

As discussed in Chapter 1, at birth, infants possess universal speech perception abilities that enable them to acquire any language. In addition, their speech perception abilities are already shaped by prenatal experience with speech, which starts at the beginning of the third trimester of gestation, when hearing becomes operational. This prenatal experience with the maternal speech already shapes newborns' listening preferences as well as how their brains respond to speech. The thesis thus asks what counts as speech for young infants: is the category already specific to spoken human language, or is it a boarder category involving primate vocalizations or communicative sounds in general? Previous literature has revealed a preference for listening to speech in newborns and infants over other auditory stimuli like white noise (Colombo & Bundi, 1981), musical instruments (Glenn, Cunningham and Joyce, 1981) or complex non-speech sine wave analogs (Vouloumanos & Werker, 2004), suggesting that speech is already special at birth. Yet, subsequent findings suggest that the category speech may be quite broad at birth, including primate vocalizations in general. Indeed, newborns do not show a listening bias for speech over monkey vocalizations behaviorally, at least when short pseudo-words are presented, while 3-months-olds do (Vouloumanos et al., 2010). We have, therefore, asked whether newborns are already sensitive to human speech when it is presented in its full complexity, or to a broader category of primate vocalizations, which then narrow down to human speech with experience?

To address this question, we conducted 4 empirical studies with newborns and young infants using a non-invasive brain imaging technique, the near-infrared spectroscopy (NIRS). We chose to use neuroimaging rather than behavioral techniques to gain more information about the underlying mechanisms, and because given the rather limited behavioral repertoire of young infants, abilities that are not robust may be seen sooner or more easily in passive tasks than when an overt behavioral response is required. We presented the basic principles and developmental applications of NIRS in some detail in Chapter 2.

Experiments 1 and 2 (Chapter 3) have measured newborns' brain responses to speech and to two non-linguistic human sounds that infants themselves can produce: newborn cries (Experiment 1) and human laughter (Experiment 2). We reasoned that finding similar brain processing mechanisms for speech and the other human sounds would argue in favor of a broad category at birth, while different neural processing mechanisms would suggest that speech is already a specific auditory stimulus for infants that triggers specialized neural mechanisms. Our results showed that the brain responses differed significantly between speech and cries/laughter. Speech triggered a left-lateralized temporal response in newborns (Experiment 1), converging with other studies, while the non-linguistic stimuli, both cries and laughter, elicited bilateral activation (Experiments 1 and 2). This pattern of results is in favor of a specific status for speech already at birth, even compared with communicative, acoustically similar human sounds that infants themselves can produce, and thus in favor of a narrow category including only speech.

In these studies, we did not measure newborns' listening preferences, but rather the brain activation in response to these different stimuli. To have a more complete picture, it would be interesting to test newborns' listening preferences for our stimuli behaviorally. Moreover, subsequent NIRS studies may explore how the newborn brain responds to monkey vocalizations as compared to speech, to explore in greater detail what has been found behaviorally (Vouloumanos et al., 2010). We hypothesize that similar processing mechanisms may be found for speech and monkey calls, when relatively short speech stimuli, e.g. single words, are used, but when more complex speech stimuli are used, e.g. full sentences with complex intonational contours and prosody, speech-specific neural mechanisms may be revealed.

Interestingly, despite their different lateralizations, speech did not trigger greater brain activation than cries or laughter, rather both newborns cries and laughter triggered greater brain activation than speech, in particular in the right hemisphere. We interpreted these findings as being linked to the emotional valence of our stimuli. While our speech stimuli were produced in an emotional-neutral way, cries and laughter are highly emotional stimuli, negative in the case of cries and positive in the case of laughter. This is in line with previous findings showing a sensitivity to emotional vocal productions compared to neutral ones already in newborns (Mastropieri & Turkewicz, 1999; Zhang et al., 2017).

Taken together, the results of Experiments 1 and 2 demonstrate that speech is already specific enough at birth to elicit specialized neural processing in a left lateralized circuitry similar to those found in most adults, compared to acoustically similar, communicative human vocal productions. Newborns can thus recognize speech as a specific auditory stimulus that requires particular attention. This ability is fundamental for language acquisition, which takes place during the next months and years. If this conclusion is true, then we can make the prediction that speech, but not other auditory stimuli, preferentially elicits specific learning mechanisms such as abstract rule learning, a foundational mechanism for learning language. To verify this prediction, more studies are needed that investigate brain learning mechanisms in newborns using non-speech stimuli as well as more naturalistic, ecologically valid, multimodal language.

To test this prediction, in Experiment 3 (Chapter 4) we investigated whether the rule learning mechanisms found to be elicited by linguistic stimuli even in newborns (Gervain et al., 2008; 2012) were specific to speech or also applied to non-linguistic stimuli. Several studies documented newborns' ability to already detect and learn structural regularities in the speech input (for a review, see de la Cruz-Pavía & Gervain, 2021). Specifically, newborns exposed to repetition-based trisyllabic pseudowords of the form AAB (e.g. "babamu") and random sequences of the form ABC (e.g. "bamuge") displayed significantly greater brain activation to the repetitive sequences AAB lateralized in the left temporal and frontal area (Gervain et al., 2012) than to the random ABC sequences. Is this sensitivity to structural regularities specific to speech? To investigate this, we exposed newborns to the same structures, i.e. AAB and ABC sequences but we replaced the syllables used in Gervain and colleagues (2012) by piano tones. Our results showed that AAB sequences also triggered stronger brain activation than ABC sequences even when presented with tones but the brain responses differed from those found in Gervain et al. (2012) in several ways: in the nature of the hemodynamic response, in the time course of the differential response, and in the localization of the activation. While syllables triggered a positive, canonical hemodynamic response, tones triggered negative, inverted responses. The inverted response to tones was due to a strong habituation effect over the time course of the study and thus a decrease of the brain response over time. By contrast, responses to syllables increased over the time course of the study, which was interpreted as learning taking place. Moreover, while the pseudowords triggered a left-lateralized response, the tone sequences triggered bilateral activation. Taken together, these results show that speech triggers specific learning mechanisms compared to some non-linguistic auditory stimuli.

Interestingly, these results parallel findings in the visual domain. Berent and colleagues (2021) tested 6-month-old infants' ability to extract repetition-based regularities from linguistic and non-linguistic stimuli in the visual modality. Infants were exposed to the repetition-based regularity in American Sign Language and in a matched non-linguistic visual control condition (cartoon drawings of a tree). While repetition-based regularities were discriminated from the random sequences in both conditions, the brain activations were different. Even more interestingly, the brain responses to signs were very similar to those found with spoken syllables in newborns (Gervain et al., 2012), despite the different sensory modality. Our results are in line with the double dissociation hypothesis raised by Berent and colleagues: linguistic stimuli, regardless of the modality in which they are produced, trigger specific brain responses, which are different from those to non-linguistic analogs, even if this latter are similar to the linguistic ones in low-level perceptual features. These results highlight the specificity of language in general for the newborn and infant brain.

Our first three experiments show the importance of early experience with speech and the role it may play for language development. But what happens when early speech experience is disrupted?

To address this question, in Chapter 5, we investigated how infants with hearing loss process speech. Specifically, we asked whether hearing loss impacts the early cortical shaping for native language processing. To do so, we employed a paradigm often used in the literature to assess the neural processing of the native language compared to an unfamiliar language and non-linguistic stimuli (backward speech) in newborns and infants (Peña et al., 2003; Dehaene et al., 2002; Sato et al., 2011; May et al., 2015; 2017; Vannasing et al., 2016; Zhang et al., 2022). Although there is some variability in the results, these studies show a left lateralized response to the native language and an advantage for, i.e. stronger response to forward as opposed to backward speech, especially in the native language. Our results reveal an early impact of hearing loss on the neural responses to languages. Indeed, infants with hearing loss did not show the typical left-lateralized activation for the native language, neither the greater activation for linguistic stimuli over non-linguistic ones (forward vs backward speech).

This study was exploratory, as our hearing loss group was relatively small and heterogeneous. Their hearing loss originated from various aetiologies, some, but not others had comorbidities, and their hearing loss varied in degree (mild to profound deafness) and in its nature (conductive hearing loss or sensorineural hearing loss). Large samples or more focused, homogeneous groups are needed to statistically assess the effect of these different factors. Moreover, our study investigated a very basic perceptual ability, the rhythmic discrimination between the native language and an unfamiliar language. Further studies could focus on more specific or more fine-grained processing abilities such as phoneme or prosodic feature discrimination in young infants with hearing loss.

The results of the HL group show that auditory deprivation/reduced auditory experience and the resulting language deprivation (if no sign language is present) have an early impact, i.e. already a few months after birth, on how the brain responds to the native language in infants. These results highlight the need for early diagnosis and intervention programs, if the family intends to oralize the child. Several possibilities are now available to at least partially restore hearing, depending on the severity of deafness. Hearing aids amplify sound and are thus recommended for patients with a mild or moderate hearing loss without lesions in the cochlea or the inner ear. For patients with profound deafness, cochlear implants (CI) are better suited. The cochlear implant is composed of different parts. First, a sound processor placed behind the ear and composed of microphones captures the environmental sounds and sends them as electrical signals to a receiver implanted under the skin, through a transmitter. Then, the receiver sends the electrical signals to electrodes implanted in the cochlea (in the inner ear), which then stimulate the auditory nerve, which transmits the signals to the brain. New technical advances have allowed very early implantation (before 12 months of age) (Naik et al., 2021) and several studies have shown that early implantation in infants with profound congenital deafness is strongly linked to better language outcomes (Purcell et al., 2021; Sharma et al., 2020 for a review). Indeed, prolonged auditory deprivation has significant impact on brain organization and it has been shown in adult CI users that stronger early auditory reorganization, i.e. the reshaping of the auditory cortex to respond to non-auditory stimuli, is linked to poorer performance in speech perception tasks after implantation (Doucet et al., 2006; Buckley & Tobey, 2011; Schierholz et al., 2015; Prince et al., 2021). Our results with deaf infants uncover the developmental origins of these effects, showing a reorganization of how the brain responds to the native language, i.e. a weak, right-lateralized pattern of activation and the absence of sensitivity for linguistic auditory stimuli. Moreover, even if bigger samples and a more focused design are needed, the activation in the temporal cortex we observed in some of our profoundly deaf infant participants may indicate an early cortical reshaping of the auditory cortex for visual processing. This very early brain reorganization thus motivates the need for early implantation in infants whose families opt for oral communication. Future studies investigating the neural processing of language in infants and children with cochlear implants are needed to better understand how CI improves speech perception as well as to more precisely investigate brain plasticity in these children.

It is very important to note that these considerations apply to patients who plan to oralize, i.e. use spoken language. Speech, i.e. the auditory modality, is of course not the only way for humans to communicate and another possibility for deaf infants is to develop sign language. In this case, auditory deprivation does not lead to linguistic deprivation and, as it has been notably shown by Berent and colleagues (2021), sign language triggers similar neural mechanisms to spoken language in 6-month-old infants. Sign language, like spoken language, displays phonological, morphological, syntactical and semantic units (Lillo-Martin, 1991). The majority of studies investigating language is based on spoken language, which does not provide a complete picture. Studying the neural processing of both spoken and signed language in parallel is a valuable way to better characterize the brain specialization for language in its abstract form. Interestingly, sign language, as spoken language, elicits a dominant activation in the left hemisphere, shedding light on what properties of language are linked to this specific left-lateralized activation, i.e. independent of surface properties (Bavelier et al., 1998; Hicock et al., 2002; MacSweeney et al., 2008).

The four experimental studies presented in this research work highlighted various interesting findings concerning early speech processing that go beyond the hypotheses we initially set out to test and reveal interesting phenomena outside of the scope of the original research questions. First, as already discussed, several of our findings confirm the early hemispheric specialization of the newborn and infant brain for the native language (Experiment 1 with French newborns, Experiment 4 with Italian-learning infants within their first year of life). This early cortical shaping has been shown by several studies in newborns (Peña et al., 2003; Sato et al., 2011; Vannasing et al., 2016), 3-month-old (Dehaene et al., 2002) and 10-month-old infants (Zhang et al., 2022) and our results strengthen these findings. Interestingly, the results of the typically hearing control group in Experiment 4 suggest a developmental change in infants' brain responses to the native language during the first year of life, meshing well with the only existing study testing rhythmic language discrimination in older infants (Zhang et al., 2022). Zhang and colleagues found bilateral and relatively weak activation to the native language in 5-month-old infants, while a left-lateralized but inverted hemodynamic response in 10-month-olds. These results suggest that the brain response to the native language may change over the first year of life. Our results from the normal-hearing infant group (1 week – 12 months) also showed a significant effect of age in the left hemisphere. Future studies will be needed to systematically test infants at various ages within the first year of life to better characterize the developmental trajectory of native language processing in the developing brain.

Second, some of our results showed that the neural mechanisms of speech processing may be influenced by context. Indeed, the same French sentences were used as the speech condition in Experiments 1 and 2, but significantly different brain responses were found between the two studies. When speech was presented in contrast with newborns cries, the typical left-lateralized canonical hemodynamic response was found. However, when the same sentences were presented in contrast with baby and adult laughter, then speech triggered inverted hemodynamic responses. In the NIRS literature, the inverted response, often found in infants, is not yet fully understood, however it has been suggested to be the reflection of the processing of unusual or degraded stimuli (Issard and Gervain, 2018). In light of this, a first possible explanation for the context effects we observed lies in the acoustic similarity between speech and the other stimuli. Speech and newborn cries are both harmonic and have similar prosodic contours (Mampe et al., 2009). By contrast, laughter is less speech-like than newborn cries, as it lacks a full acoustic contour. These differences may trigger different brain responses to speech as a function of the acoustic distance between the stimuli presented together. Another, not necessarily mutually exclusive possibility is the influence of the emotional valence of the stimuli. In Experiment 2, in which speech triggered an inverted response, two conditions of adult vocalizations were presented: spoken sentences and laughter. The spoken sentences were produced with no particular emotional valence, i.e. in a neutral way, while laughter is a highly emotional, positive stimulus. The strong activation in response to adult laughter and the inverted response to speech may reflect a listening bias for emotional adult vocalizations over neutral ones, as caregivers' positive affect may be a particularly important cue for newborns' care, well-being and survival. One interesting future study would be to test whether replacing the neutral sentences by sentences produced with a happy prosody would change how the newborn brain processes speech compared with laughter.

Third, our work may provide new insights on the links between the production and perception of speech. The motor theory of speech perception (Liberman, 1967) proposes that there is a key link between speech perception and speech production, arguing that the perception of speech is based on the retrieval of the articulatory gestures that enable the production of the sounds we hear when we produce them. This theory has rarely been studied in developmental populations, but some recent findings suggest a sensorimotor influence on speech perception in 6-month-old infants, while several studies seem to indicate that infants prefer to listen to stimuli that are more similar to what they can themselves produce (DePaolis et al., 2011; 2013; Polka et al., 2022). In Experiment 1, we investigated newborns' brain responses to stimuli that they themselves can produce and use for communication: newborn cries, while in Experiment 2, we used two conditions of laughter, baby laughter and adult laughter, i.e. one type of laughter infants can produce and another they cannot, in order to search for specific brain responses to stimuli that are close to newborns' vocal abilities. Newborn cries triggered a significant bilateral activation, greater than the activation in response to speech. Baby laughter triggered similar bilateral activation, which did not differ from that triggered by cries. These results may suggest a shared brain network activated by "producible" stimuli in the infant brain. Imada and colleagues (2006) investigated the activation of the speech motor analysis area during speech perception in neonates, 6- and 12-month-old infants. They found that the premotor brain area was activated when 6- and 12-month-old infants, but not newborns, were listening to speech. They hypothesized that premotor activation is linked to perceiving producible linguistic stimuli. Our results converge with these findings and extend them to non-linguistic communicative sounds. Future studies could specifically target the premotor area in newborns when they are exposed to stimuli they produce: newborn cries. This would provide new insights on the emergence of the perceptual-motor link in auditory perception.

In this research work, we have investigated speech processing in newborns and young infants. Our results have highlighted the specific status of speech for the newborn and the infant brain compared to other auditory communicative and non-linguistic stimuli. How do these results relate to what we know more generally about language development? At birth, newborns show universal speech perception abilities that allow them to acquire any language, even languages that are different from the one(s) heard prenatally, as multiple studies with adopted children highlighted (Pierce et al., 2014; Ventureyra et al., 2004). In addition to these universal abilities, newborns already show effects of their prenatal exposure to language. This prenatal experience with speech shapes newborns' perception abilities: newborns are already sensitive to the prosodic features of their native language (Abboub et al., 2016) and can already recognize their native language based on rhythm (Mehler et al., 1988), which triggered a specific left-lateralized activation (Experiments 1 and 4). Indeed, melody and rhythm is how infants first discover speech: the first experience with speech starts prenatally, when speech is low-pass filtered by the uterine tissues, thus only preserving the prosody of speech, i.e. melody and rhythm. A new, exciting hypothesis (Gervain, 2015; Nallet & Gervain, 2021) suggests that prosody is foundational for language development and links the innate predispositions of infants for language learning to language experience during development (Figure 6.1). At the heart of this hypothesis lies the theory of embedded neural oscillations, an influential model of speech perception, originally proposed for adults (Ghitza 2011, Giraud & Poeppel 2012). This model emphasizes the close correspondence between the various linguistic units appearing at different time scales in speech (Gervain 2018, Nespor & Vogel 1986, Selkirk 1986) and brain oscillations operating at different frequency bands. By matching the specific temporal scales of the different linguistic units, the neural oscillations parse the incoming speech signal into its constituent units simultaneously at different levels, from phonemes through syllables to larger phrases, thus allowing speech and language processing. Three frequency bands contribute to the bottom-up processing of the speech signal: delta oscillations (1–3 Hz) match prosodic phrases, theta oscillations (4–8 Hz) correspond to syllables, and low-gamma oscillations (>35 Hz) match (sub)phonemic units (Giraud & Poeppel 2012). These different brain oscillations are hierarchically nested, i.e., the faster oscillations are embedded in the slower ones. This phase-amplitude coupling between the different frequency bands (Buzsáki 2006) has been linked to cognitive performance and speech perception (Goswami 2020, Ladányi et al. 2020, Nallet &

Gervain, 2022). Additionally, the beta band (12–25 Hz) has more recently been shown to contribute top-down information, that is, predictions for upcoming linguistic content on the basis of the listener's linguistic knowledge (Fontolan et al. 2014, Hovsepyan et al. 2020). The NIRS measurements we used in this thesis did not allow us to record the activity of the brain oscillations in the auditory cortex during speech listening. Thus, it would be very interesting in future studies to put together our findings with electroencephalographic measures (EEG) in response to the same stimuli.

Linking what we know about newborns' speech perception abilities and this neural model of speech perception, we speculate that the oscillatory hierarchy emerges from a unique developmental chronology of infants' language experience, which starts during prenatal life. During the last trimester of pregnancy, fetuses' experience with low-pass filtered speech, which mainly preserves prosody, contributes to the attunement of the slower oscillations, delta and theta, to the prosodic properties of the native language. After birth, infants start to experience the full-band speech signal that contains the finer acoustic details necessary for the discrimination of speech sounds present in the native language. After several months of exposure to this complete speech signal, gamma oscillations, responsible for phoneme perception, are fine-tuned and get embedded in the slower, already native-like delta and theta oscillations. The developmental chronology of experience with language—first the filtered prenatal signal, followed by the full-band postnatal signal—corresponds to the oscillatory hierarchy in the auditory cortex. Our hypothesis is that prosody, the highest element of the hierarchy in which the others are embedded and the first linguistic information available to infants, guides speech perception and language acquisition. Indeed, the bootstrapping role of prosody in older infants is well established (Jusczyk 2001, Morgan & Demuth 1996, Nespor et al. 2008). Babies rely on prosodic cues to segment words (Shukla et al. 2011) or to determine the basic word order of their native language (Gervain & Werker 2013).

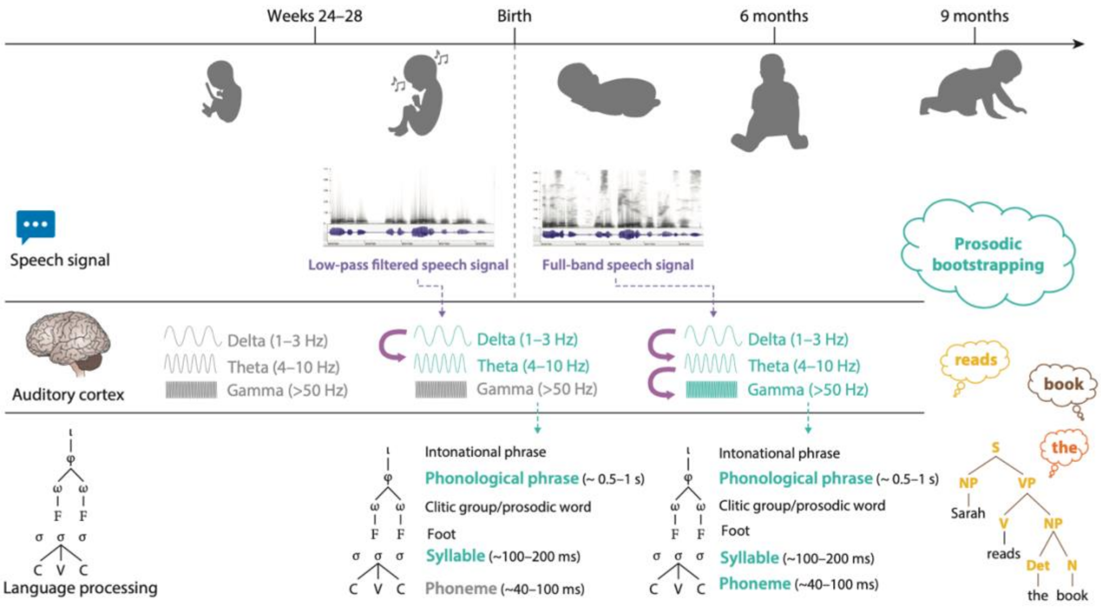


Figure 6.1 : Schematic illustration of the prenatal prosodic shaping hypothesis proposing that the speech perception system, supported by neural oscillations in the auditory cortex, develops through the chronology of infants' experience with speech, starting prenatally with prosody as the crucial cue guiding speech perception and further language development. Abbreviations: C, consonant; Det, determiner; N, noun; NP, noun phrase; S, sentence; V, verb (in the syntactic tree); V, vowel (in the linguistic trees); VP, verb phrase. (From Nallet & Gervain, 2021)

This hypothesis thus suggests that the prenatal period plays a fundamental role in the typical development of the speech perception system. One fundamental study for this hypothesis (Ortiz Barajas et al. 2021) testing newborns' and 6-month-olds' ability to track the speech envelope (i.e., match theta oscillations to the syllable rate) indeed suggests that neural oscillations in response to language are already operational at birth. Further studies are necessary to assess the presence and functions of the entire oscillatory cascade in newborns and young infants and to test their specificity for speech and language. If empirically confirmed, this hypothesis could provide important insights into language acquisition, and it may also have important implications for atypical language development, as some children do not have a typical prenatal experience (e.g., due to preterm birth or congenital deafness).

7 Conclusion

In the recent decades, great advances have been made in our understanding of how humans acquire language. Results have shown that learning begins very early, as evidenced by newborns' remarkable speech perception abilities. At birth, infants show impressive universal, broad-based speech perception abilities, making them "linguistic citizens of the world", able to learn and acquire any language. Impressively, newborns do not represent the "initial state" as language discovery starts even before birth. The journey to language acquisition starts at the beginning of the third trimester of pregnancy during which fetuses discover speech as a distant, but intriguing, melody. Indeed, native language prosody, i.e. its melody and rhythm, preserved even in the intrauterine environment, is the first speech cue infants perceive prenatally. The development of brain imaging techniques and their applications in newborns have revealed a neurodevelopmental preparedness for language in the neonatal brain.

The current work, by investigating the neural processing mechanisms of speech in newborns and young infants, has provided new insights into early language perception and its correlated brain mechanisms. Specifically, it has brought new findings highlighting the privileged status of speech among other human vocal productions for newborns. We have shown that speech may be salient enough for newborns to elicit specialized brain processing, different from that elicited by other human, communicative and acoustically similar sounds. Furthermore, we provided new evidence for the existence of specific neural learning mechanisms in newborns for language, regardless of the modality, suggesting a special status of human language for infant perception. This special status and the particular attention and brain processing it entails provides a better understanding of the surprisingly fast and early learning abilities infants and young children demonstrate. Finally, we highlighted that atypical speech experience induced by hearing loss may precociously affect how the infant brain processes speech, suggesting the need for early diagnosis and intervention in case parents decide to oralize their child, to minimize the detrimental impact of hearing loss on language acquisition, academic achievement and general well-being of children with hearing loss.

This research work, along with the numerous studies investigating the newborn and young infant's brain processing of speech, sheds light on the impressive plasticity of the brain in the beginning of life. We have seen that the brain responses to auditory stimuli is highly linked to auditory experience: the newborn brain activation is already shaped by what the baby has heard in utero, while a reduced auditory stimulation has an extremely early impact on the functioning of the auditory cortex.

References

Abboub, N., Nazzi, T. and Gervain, J. (2016) 'Prosodic grouping at birth', *Brain and Language*, 162, pp. 46–59. Available at: https://doi.org/10.1016/j.bandl.2016.08.002.

Abrams, R.M. and Gerhardt, K.J. (2000) 'The Acoustic Environment and Physiological Responses of the Fetus', *Journal of Perinatology*, 20(S1), pp. S31–S36. Available at: https://doi.org/10.1038/sj.jp.7200445.

Akinpelu, O.V. *et al.* (2014) 'Otoacoustic emissions in newborn hearing screening: A systematic review of the effects of different protocols on test outcomes', *International Journal of Pediatric Otorhinolaryngology*, 78(5), pp. 711–717. Available at: https://doi.org/10.1016/j.ijporl.2014.01.021.

Arpino, C. *et al.* (2010) 'Preterm birth and neurodevelopmental outcome: a review', *Child's Nervous System*, 26(9), pp. 1139–1149. Available at: https://doi.org/10.1007/s00381-010-1125-y.

Aslin, R.N. and Mehler, J. (2005) 'Near-infrared spectroscopy for functional studies of brain activity in human infants: promise, prospects, and challenges', *Journal of Biomedical Optics*, 10(1), p. 011009. Available at: https://doi.org/10.1117/1.1854672.

Barre, N. *et al.* (2011) 'Language Abilities in Children Who Were Very Preterm and/or Very Low Birth Weight: A Meta-Analysis', *The Journal of Pediatrics*, 158(5), pp. 766-774.e1. Available at: https://doi.org/10.1016/j.jpeds.2010.10.032.

Bartha-Doering, L. *et al.* (2019) 'Absence of neural speech discrimination in preterm infants at term-equivalent age', *Developmental Cognitive Neuroscience*, 39, p. 100679. Available at: https://doi.org/10.1016/j.dcn.2019.100679.

Bavelier, D., Corina, D. P., & Neville, H. J. (1998). Brain and language: a perspective from sign language. *Neuron*, *21*(2), 275-278.

Benavides-Varela, S. and Gervain, J. (2017) 'Learning word order at birth: A NIRS study', *Developmental Cognitive Neuroscience*, 25, pp. 198–208. Available at: https://doi.org/10.1016/j.dcn.2017.03.003.

Berent, I. *et al.* (2021) 'Infants differentially extract rules from language', *Scientific Reports*, 11(1), p. 20001. Available at: https://doi.org/10.1038/s41598-021-99539-8.

Bijeljac-Babic, R., Bertoncini, J. and Mehler, J. (1993) 'How do 4-day-old infants categorize multisyllabic utterances?', *Developmental Psychology*, 29(4), pp. 711–721. Available at: https://doi.org/10.1037/0012-1649.29.4.711.

Binder, J.R. *et al.* (1995) 'Lateralized Human Brain Language Systems Demonstrated by Task Subtraction Functional Magnetic Resonance Imaging', *Archives of*

Neurology, 52(6), pp. 593–601. Available at: https://doi.org/10.1001/archneur.1995.00540300067015.

Biswas, A.K. *et al.* (2012) ‘The Potential Risk Factors and the Identification of Hearing Loss in Infants’, *Indian Journal of Otolaryngology and Head & Neck Surgery*, 64(3), pp. 214–217. Available at: https://doi.org/10.1007/s12070-011-0307-6.

Blasi, A. *et al.* (2011) ‘Early Specialization for Voice and Emotion Processing in the Infant Brain’, *Current Biology*, 21(14), pp. 1220–1224. Available at: https://doi.org/10.1016/j.cub.2011.06.009.

Boersma, P. and van Heuven, V. (2001) ‘Speak and unSpeak with PRAAT’, *Glot International*, 5.

Bortolato, T. *et al.* (2023) ‘Slow development of vocal sequences through ontogeny in wild chimpanzees (*Pan troglodytes verus)*’, *Developmental Science*, 26(4), p. e13350. Available at: https://doi.org/10.1111/desc.13350.

Bouchon, C., Abboub, N., Gervain, J. (2013). NIRS (Near Infrared Spectroscopy): Principes généraux de cette nouvelle technique de neuroimagerie et contribution à la recherche aux origins du langage. *Les Cahiers de l’Audition*. N°3/2013.

Bouchon, C., Nazzi, T. and Gervain, J. (2015) ‘Hemispheric Asymmetries in Repetition Enhancement and Suppression Effects in the Newborn Brain’, *PLOS ONE*. Edited by D.S. Vicario, 10(10), p. e0140160. Available at: https://doi.org/10.1371/journal.pone.0140160.

Broca, P.M. (1861). Remarques sur le siége de la faculté du langage articulé, suivies d’une observation d’aphémie (perte de la parole). *Bull. Soc. Anat.* **6**, 330–357.

Bruderer, A.G. *et al.* (2015) ‘Sensorimotor influences on speech perception in infancy’, *Proceedings of the National Academy of Sciences*, 112(44), pp. 13531–13536. Available at: https://doi.org/10.1073/pnas.1508631112.

Bryant, G.A. and Aktipis, C.A. (2014) ‘The animal nature of spontaneous human laughter’, *Evolution and Human Behavior*, 35(4), pp. 327–335. Available at: https://doi.org/10.1016/j.evolhumbehav.2014.03.003.

Busnel, M-C., Granier-Deferre, C. (1983). And what of fetal audition? In *The Behavior of Human Infants*, ed. A Oliverio, M Zappella, pp. 93-126. Boston, MA: Springer

Byers-Heinlein, K., Burns, T.C. and Werker, J.F. (2010) ‘The Roots of Bilingualism in Newborns’, *Psychological Science*, 21(3), pp. 343–348. Available at: https://doi.org/10.1177/0956797609360758.

Cabrera, L. and Gervain, J. (2020) ‘Speech perception at birth: The brain encodes fast and slow temporal information’, *Science Advances*, 6(30), p. eaba7830. Available at: https://doi.org/10.1126/sciadv.aba7830.

Cheour, M. *et al.* (2002) ‘Speech sounds learned by sleeping newborns’, *Nature*, 415(6872), pp. 599–600. Available at: https://doi.org/10.1038/415599b.

Choi, D., Black, A.K. and Werker, J.F. (2018) 'Cascading and Multisensory Influences on Speech Perception Development', *Mind, Brain, and Education*, 12(4), pp. 212–223. Available at: https://doi.org/10.1111/mbe.12162.

Chomsky, N. (1959) 'On certain formal properties of grammars', *Information and Control*, 2(2), pp. 137–167. Available at: https://doi.org/10.1016/S0019-9958(59)90362-6.

Christophe, A. *et al.* (1994) 'Do infants perceive word boundaries? An empirical study of the bootstrapping of lexical acquisition', *The Journal of the Acoustical Society of America*, 95(3), pp. 1570–1580. Available at: https://doi.org/10.1121/1.408544.

Claeskens, G., & Hjort, N. L. (2008). Model selection and model averaging. *Cambridge books*. Cambridge University Press.

Colombo, J. and Bundy, R.S. (1981) 'A method for the measurement of infant auditory selectivity', *Infant Behavior and Development*, 4, pp. 219–223. Available at: https://doi.org/10.1016/S0163-6383(81)80025-2.

Coye, C., Zuberbühler, K. and Lemasson, A. (2016) 'Morphologically structured vocalizations in female Diana monkeys', *Animal Behaviour*, 115, pp. 97–105. Available at: https://doi.org/10.1016/j.anbehav.2016.03.010.

Davila-Ross, M. *et al.* (2011) 'Aping expressions? Chimpanzees produce distinct laugh types when responding to laughter of others.', *Emotion*, 11(5), pp. 1013–1020. Available at: https://doi.org/10.1037/a0022594.

Dawson, C. and Gerken, L. (2009) 'From domain-generality to domain-sensitivity: 4-Month-olds learn an abstract repetition rule in music that 7-month-olds do not', *Cognition*, 111(3), pp. 378–382. Available at: https://doi.org/10.1016/j.cognition.2009.02.010.

De Haan, M. (ed.) (2013) *Infant EEG and Event-Related Potentials*. 0 edn. Psychology Press. Available at: https://doi.org/10.4324/9780203759660.

De La Cruz-Pavía, I. and Gervain, J. (2021) 'Infants' perception of repetition-based regularities in speech: a look from the perspective of the same/different distinction', *Current Opinion in Behavioral Sciences*, 37, pp. 125–132. Available at: https://doi.org/10.1016/j.cobeha.2020.11.014.

De La Cruz-Pavía, I. and Gervain, J. (2023) 'Six-month-old infants' perception of structural regularities in speech', *Cognition*, 238, p. 105526. Available at: https://doi.org/10.1016/j.cognition.2023.105526.

DeCasper, A.J. *et al.* (1994) 'Fetal reactions to recurrent maternal speech', *Infant Behavior and Development*, 17(2), pp. 159–164. Available at: https://doi.org/10.1016/0163-6383(94)90051-5.

DeCasper, A.J. and Fifer, W.P. (1980) 'Of Human Bonding: Newborns Prefer Their Mothers' Voices', *Science*, 208(4448), pp. 1174–1176. Available at: https://doi.org/10.1126/science.7375928.

DeCasper, A.J. and Sigafoos, A.D. (1983) ‘The intrauterine heartbeat: A potent reinforcer for newborns’, *Infant Behavior and Development*, 6(1), pp. 19–25. Available at: https://doi.org/10.1016/S0163-6383(83)80004-6.

DeCasper, A.J. and Spence, M.J. (1986) ‘Prenatal maternal speech influences newborns’ perception of speech sounds’, *Infant Behavior and Development*, 9(2), pp. 133–150. Available at: https://doi.org/10.1016/0163-6383(86)90025-1.

Dehaene, S. *et al.* (1997) ‘Anatomical variability in the cortical representation of first and second language’:, *NeuroReport*, 8(17), pp. 3809–3815. Available at: https://doi.org/10.1097/00001756-199712010-00030.

Dehaene-Lambertz, G. and Baillet, S. (1998) ‘A phonological representation in the infant brain’:, *NeuroReport*, 9(8), pp. 1885–1888. Available at: https://doi.org/10.1097/00001756-199806010-00040.

Dehaene-Lambertz, G., Dehaene, S. and Hertz-Pannier, L. (2002) ‘Functional Neuroimaging of Speech Perception in Infants’, *Science*, 298(5600), pp. 2013–2015. Available at: https://doi.org/10.1126/science.1077066.

Dehaene-Lambertz, G. and Pena, M. (2001) ‘Electrophysiological evidence for automatic phonetic processing in neonates’:, *Neuroreport*, 12(14), pp. 3155–3158. Available at: https://doi.org/10.1097/00001756-200110080-00034.

Della Longa, L. *et al.* (2021) ‘Psychophysiological and Visual Behavioral Responses to Faces Associated with Affective and Non-affective Touch in Four-month-old Infants’, *Neuroscience*, 464, pp. 67–78. Available at: https://doi.org/10.1016/j.neuroscience.2020.07.053.

Dellwo, V. (2006) ‘Rhythm and Speech Rate: A Variation Coefficient for deltaC’. Available at: https://doi.org/10.5167/UZH-111789.

DePaolis, R.A., Vihman, M.M. and Keren-Portnoy, T. (2011) ‘Do production patterns influence the processing of speech in prelinguistic infants?’, *Infant Behavior and Development*, 34(4), pp. 590–601. Available at: https://doi.org/10.1016/j.infbeh.2011.06.005.

DePaolis, R.A., Vihman, M.M. and Nakai, S. (2013) ‘The influence of babbling patterns on the processing of speech’, *Infant Behavior and Development*, 36(4), pp. 642–649. Available at: https://doi.org/10.1016/j.infbeh.2013.06.007.

Di Giorgio, E. *et al.* (2012) ‘Is the face-perception system human-specific at birth?’, *Developmental Psychology*, 48(4), pp. 1083–1090. Available at: https://doi.org/10.1037/a0026521.

Dondi, M., Simion, F. and Caltran, G. (1999) ‘Can newborns discriminate between their own cry and the cry of another newborn infant?’, *Developmental Psychology*, 35(2), pp. 418–426. Available at: https://doi.org/10.1037/0012-1649.35.2.418.

Dubois, J. *et al.* (2014) ‘The early development of brain white matter: A review of imaging studies in fetuses, newborns and infants’, *Neuroscience*, 276, pp. 48–71. Available at: https://doi.org/10.1016/j.neuroscience.2013.12.044.

Eggermont, J.J., Moore, J.K. (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.

Eimas, P.D. *et al.* (1971) ‘Speech Perception in Infants’, *Science*, 171(3968), pp. 303–306. Available at: https://doi.org/10.1126/science.171.3968.303.

Elman, J., Bates, E., Johnson, M., Karmiloff-Smith, A., Parisi, D., Plunkett, K. (1997). Rethinking Innateness: A Connectionist Perspective on Development. Boston: MIT Press

Endress, A.D., Nespor, M. and Mehler, J. (2009) ‘Perceptual and memory constraints on language acquisition’, *Trends in Cognitive Sciences*, 13(8), pp. 348–353. Available at: https://doi.org/10.1016/j.tics.2009.05.005.

Ethofer, T. *et al.* (2007) ‘The voices of seduction: cross-gender effects in processing of erotic prosody’, *Social Cognitive and Affective Neuroscience*, 2(4), pp. 334–337. Available at: https://doi.org/10.1093/scan/nsm028.

Fantz, R.L. (1961) ‘The Origin of Form Perception’, *Scientific American*, 204(5), pp. 66–73. Available at: https://doi.org/10.1038/scientificamerican0561-66.

Ferguson, B., Franconeri, S.L. and Waxman, S.R. (2018) ‘Very young infants learn abstract rules in the visual modality’, *PLOS ONE*. Edited by M.S. Malmierca, 13(1), p. e0190185. Available at: https://doi.org/10.1371/journal.pone.0190185.

Ferreri, L. *et al.* (2014) ‘The promise of Near-Infrared Spectroscopy (NIRS) for psychological research: A brief review’:, *L’Année psychologique*, Vol. 114(3), pp. 537–569. Available at: https://doi.org/10.3917/anpsy.143.0537.

Fifer, W.P. *et al.* (2010) ‘Newborn infants learn during sleep’, *Proceedings of the National Academy of Sciences*, 107(22), pp. 10320–10323. Available at: https://doi.org/10.1073/pnas.1005061107.

Floccia, C., Christophe, A. and Bertoncini, J. (1997) ‘High-Amplitude Sucking and Newborns: The Quest for Underlying Mechanisms’, *Journal of Experimental Child Psychology*, 64(2), pp. 175–198. Available at: https://doi.org/10.1006/jecp.1996.2349.

Flom, R. and Bahrick, L.E. (2007) ‘The development of infant discrimination of affect in multimodal and unimodal stimulation: The role of intersensory redundancy.’, *Developmental Psychology*, 43(1), pp. 238–252. Available at: https://doi.org/10.1037/0012-1649.43.1.238.

Fontolan L, Morillon B, Liegeois-Chauvel C, et al. (2014). The contribution of frequency-specific activity to hierarchical information processing in the human auditory cortex. *Nat. Commun.* 5:4694

Forgács, B. *et al.* (2022) 'The newborn brain is sensitive to the communicative function of language', *Scientific Reports*, 12(1), p. 1220. Available at: https://doi.org/10.1038/s41598-022-05122-0.

Fortnum, H. and Davis, A. (1997) 'Epidemiology of Permanent Childhood Hearing Impairment in Trent Region, 1985–1993', *British Journal of Audiology*, 31(6), pp. 409–446. Available at: https://doi.org/10.3109/03005364000000037.

Friederici, A.D. (2012) 'The cortical language circuit: from auditory perception to sentence comprehension', *Trends in Cognitive Sciences*, 16(5), pp. 262–268. Available at: https://doi.org/10.1016/j.tics.2012.04.001.

Geangu, E. *et al.* (2010) 'Contagious crying beyond the first days of life', *Infant Behavior and Development*, 33(3), pp. 279–288. Available at: https://doi.org/10.1016/j.infbeh.2010.03.004.

Gemignani, J. *et al.* (2023) 'Reproducibility of infant fNIRS studies: a meta-analytic approach', *Neurophotonics*, 10(02). Available at: https://doi.org/10.1117/1.NPh.10.2.023518.

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

Gerhardt, K.J. *et al.* (1992) 'Cochlear microphonics recorded from fetal and newborn sheep', *American Journal of Otolaryngology*, 13(4), pp. 226–233. Available at: https://doi.org/10.1016/0196-0709(92)90026-P.

Gervain, J. *et al.* (2008) 'The neonate brain detects speech structure', *Proceedings of the National Academy of Sciences*, 105(37), pp. 14222–14227. Available at: https://doi.org/10.1073/pnas.0806530105.

Gervain, J. *et al.* (2011) 'Near-infrared spectroscopy: A report from the McDonnell infant methodology consortium', *Developmental Cognitive Neuroscience*, 1(1), pp. 22–46. Available at: https://doi.org/10.1016/j.dcn.2010.07.004.

Gervain, J. (2015) 'Plasticity in early language acquisition: the effects of prenatal and early childhood experience', *Current Opinion in Neurobiology*, 35, pp. 13–20. Available at: https://doi.org/10.1016/j.conb.2015.05.004.

Gervain, J., Berent, I. and Werker, J.F. (2012) 'Binding at Birth: The Newborn Brain Detects Identity Relations and Sequential Position in Speech', *Journal of Cognitive Neuroscience*, 24(3), pp. 564–574. Available at: https://doi.org/10.1162/jocn_a_00157.

Gervain, J. and Mehler, J. (2010) ‘Speech Perception and Language Acquisition in the First Year of Life’, *Annual Review of Psychology*, 61(1), pp. 191–218. Available at: https://doi.org/10.1146/annurev.psych.093008.100408.

Gervain J, Werker JF. (2013). Prosody cues word order in 7-month-old bilingual infants. *Nat. Commun.* 4(1):1490

Gervain J. (2018). Gateway to language: the perception of prosody at birth. In *Boundaries Crossed, at the Interfaces of Morphosyntax, Phonology, Pragmatics and Semantics*, ed. H Bartos, M den Dikken, Z Bánréti, T Váradi, pp. 373–84. Cham, Switz.: Springer

Gervais, M. and Wilson, D.S. (2005) ‘THE EVOLUTION AND FUNCTIONS OF LAUGHTER AND HUMOR: A SYNTHETIC APPROACH’, *The Quarterly Review of Biology*, 80(4), pp. 395–430. Available at: https://doi.org/10.1086/498281.

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

Ginzburg, J., Breitholtz, E., Cooper, R., Hough, J., Tian, Y. (2015). Understanding Laughter. 20th Amsterdam Colloquuium, Amsterdam, Netherlands.

Ginzburg, J., Mazzocconi, C. and Tian, Y. (2020) ‘Laughter as language’, *Glossa: a journal of general linguistics*, 5(1). Available at: https://doi.org/10.5334/gjgl.1152.

Giraud A-L, Poeppel D. (2012). Cortical oscillations and speech processing: emerging computational principles and operations. *Nat. Neurosci.* 15(4):511–17

Glenn, S.M., Cunningham, C.C. and Joyce, P.F. (1981) ‘A Study of Auditory Preferences in Nonhandicapped Infants and Infants with down’s Syndrome’, *Child Development*, 52(4), p. 1303. Available at: https://doi.org/10.2307/1129520.

Gómez, D.M. *et al.* (2014) ‘Language universals at birth’, *Proceedings of the National Academy of Sciences*, 111(16), pp. 5837–5841. Available at: https://doi.org/10.1073/pnas.1318261111.

Gonzalez-Gomez, N. and Nazzi, T. (2012) ‘Phonotactic acquisition in healthy preterm infants’, *Developmental Science*, 15(6), pp. 885–894. Available at: https://doi.org/10.1111/j.1467-7687.2012.01186.x.

Goren, C.C., Sarty, M. and Wu, P.Y.K. (1975) ‘Visual Following and Pattern Discrimination of Face-like Stimuli by Newborn Infants’, *Pediatrics*, 56(4), pp. 544–549. Available at: https://doi.org/10.1542/peds.56.4.544.

Goswami U. (2020). Toward realizing the promise of educational neuroscience: improving experimental design in developmental cognitive neuroscience studies. *Annu. Rev. Dev. Psychol.* 2:133–55

Grabe, E. and Low, E.L. (2002) ‘Durational variability in speech and the Rhythm Class Hypothesis’, in C. Gussenhoven and N. Warner (eds) *Laboratory Phonology 7*. Mouton de Gruyter, pp. 515–546. Available at: https://doi.org/10.1515/9783110197105.2.515.

Granier-Deferre, C. *et al.* (2011) ‘A Melodic Contour Repeatedly Experienced by Human Near-Term Fetuses Elicits a Profound Cardiac Reaction One Month after Birth’, *PLoS ONE*. Edited by G. Chapouthier, 6(2), p. e17304. Available at: https://doi.org/10.1371/journal.pone.0017304.

Grantham, H. *et al.* (2022) ‘Effects of Segmental and Suprasegmental Speech Perception on Reading in Pediatric Cochlear Implant Recipients’, *Journal of Speech, Language, and Hearing Research*, 65(9), pp. 3583–3594. Available at: https://doi.org/10.1044/2022_JSLHR-22-00035.

Griffiths, S.K. *et al.* (1994) ‘The perception of speech sounds recorded within the uterus of a pregnant sheep’, *The Journal of the Acoustical Society of America*, 96(4), pp. 2055–2063. Available at: https://doi.org/10.1121/1.410147.

Grossmann, T. *et al.* (2010) ‘The Developmental Origins of Voice Processing in the Human Brain’, *Neuron*, 65(6), pp. 852–858. Available at: https://doi.org/10.1016/j.neuron.2010.03.001.

Grossmann, T., (2013). The early development of processing emotions in face and voice. In: Belin, P., Campanella, S., Ethofer, T. (Eds.), *Integrating Face and Voice in Person Perception.* Springer New York, New York, NY, pp. 95–116. http://dx.doi.org/10. 1007/978-1-4614-3585-3_5. Retrieved from.

Guarini, A. *et al.* (2009) ‘Reconsidering the impact of preterm birth on language outcome’, *Early Human Development*, 85(10), pp. 639–645. Available at: https://doi.org/10.1016/j.earlhumdev.2009.08.061.

Guarini, A. *et al.* (2016) ‘Linguistic features in children born very preterm at preschool age’, *Developmental Medicine & Child Neurology*, 58(9), pp. 949–956. Available at: https://doi.org/10.1111/dmcn.13118.

Gustafson, G.E. *et al.* (2017) ‘Newborns’ Cries are Unique to Individuals (But Not to Language Environment)’, *Infancy*, 22(6), pp. 736–747. Available at: https://doi.org/10.1111/infa.12192.

Henson, R.N.A. and Rugg, M.D. (2003) ‘Neural response suppression, haemodynamic repetition effects, and behavioural priming’, *Neuropsychologia*, 41(3), pp. 263–270. Available at: https://doi.org/10.1016/S0028-3932(02)00159-8.

Herold, B. *et al.* (2008) ‘Impaired word stress pattern discrimination in very-low-birthweight infants during the first 6 months of life’, *Developmental Medicine & Child Neurology*, 50(9), pp. 678–683. Available at: https://doi.org/10.1111/j.1469-8749.2008.03055.x.

Hickok, G. and Poeppel, D. (2007) 'The cortical organization of speech processing', *Nature Reviews Neuroscience*, 8(5), pp. 393–402. Available at: https://doi.org/10.1038/nrn2113.

Holt, C.M., Demuth, K. and Yuen, I. (2016) 'The Use of Prosodic Cues in Sentence Processing by Prelingually Deaf Users of Cochlear Implants', *Ear & Hearing*, 37(4), pp. e256–e262. Available at: https://doi.org/10.1097/AUD.0000000000000253.

Homae, F. *et al.* (2006) 'The right hemisphere of sleeping infant perceives sentential prosody', *Neuroscience Research*, 54(4), pp. 276–280. Available at: https://doi.org/10.1016/j.neures.2005.12.006.

Hovsepyan S, Olasagasti I, Giraud AL. (2020). Combining predictive coding and neural oscillations enables online syllable recognition in natural speech. *Nat. Commun.* 11:3117

Huotilainen, M. *et al.* (2005) 'Short-term memory functions of the human fetus recorded with magnetoencephalography':, *NeuroReport*, 16(1), pp. 81–84. Available at: https://doi.org/10.1097/00001756-200501190-00019.

Imada, T. *et al.* (2006) 'Infant speech perception activates Broca's area: a developmental magnetoencephalography study', *NeuroReport*, 17(10), pp. 957–962. Available at: https://doi.org/10.1097/01.wnr.0000223387.51704.89.

Issard, C. and Gervain, J. (2018) 'Variability of the hemodynamic response in infants: Influence of experimental design and stimulus complexity', *Developmental Cognitive Neuroscience*, 33, pp. 182–193. Available at: https://doi.org/10.1016/j.dcn.2018.01.009.

Izard, V. *et al.* (2009) 'Newborn infants perceive abstract numbers', *Proceedings of the National Academy of Sciences*, 106(25), pp. 10382–10385. Available at: https://doi.org/10.1073/pnas.0812142106.

James, A.R. (1988) 'Marina Nespor & Irene Vogel (1986). Prosodic phonology. Dordrecht: Foris Publications. Pp. xiv + 327.', *Phonology*, 5(1), pp. 161–168. Available at: https://doi.org/10.1017/S0952675700002219.

Jardri, R. *et al.* (2008) 'Fetal cortical activation to sound at 33 weeks of gestation: A functional MRI study', *NeuroImage*, 42(1), pp. 10–18. Available at: https://doi.org/10.1016/j.neuroimage.2008.04.247.

Johnson, M.H. *et al.* (1991) 'Newborns' preferential tracking of face-like stimuli and its subsequent decline', *Cognition*, 40(1–2), pp. 1–19. Available at: https://doi.org/10.1016/0010-0277(91)90045-6.

Johnson, M. H., & Morton, J. (1991). *Biology and cognitive development: The case of face recognition.* Oxford, England: Blackwell.

Jusczyk PW. (2001). Bootstrapping from the signal: some further directions. In *Language Acquisition and Language Disorders*, ed. J Weissenborn, B Höhle, pp. 3–23. Amsterdam: John Benjamins

Kalathottukaren, R.T., Purdy, S.C. and Ballard, E. (2017) 'Prosody Perception and Production in Children with Hearing Loss and Age- and Gender-Matched Controls', *Journal of the American Academy of Audiology*, 28(04), pp. 283–294. Available at: https://doi.org/10.3766/jaaa.16001.

Kuhl, P.K. (1986). Theoretical contributions of tests on animals to the special-mechanisms debate in speech. *Exp. Biol*. 45 :233-65

Kelsey, C. *et al.* (2023) 'Shedding light on functional near-infrared spectroscopy and open science practices', *Neurophotonics*, 10(02). Available at: https://doi.org/10.1117/1.NPh.10.2.023520.

Kennedy, C.R. *et al.* (2006) 'Language Ability after Early Detection of Permanent Childhood Hearing Impairment', *New England Journal of Medicine*, 354(20), pp. 2131–2141. Available at: https://doi.org/10.1056/NEJMoa054915.

Kitamura, C., & Burnham, D. (1998). The infant's response to maternal vocal affect. In C. Rovee-Collier, L. Lipsitt, & H. Hayne (Eds.), *Advances in infancy research* (Vol. 12, pp. 221–236). Stamford, CT: Ablex.

Knecht, S. (2000) 'Handedness and hemispheric language dominance in healthy humans', *Brain*, 123(12), pp. 2512–2518. Available at: https://doi.org/10.1093/brain/123.12.2512.

Konadath, S., Raveendran, R. and Yeshoda, K. (2021) 'Perception of speech stress in children with hearing impairment', *International Journal of Pediatric Otorhinolaryngology*, 140, p. 110495. Available at: https://doi.org/10.1016/j.ijporl.2020.110495.

Kotilahti, K. *et al.* (2010) 'Hemodynamic responses to speech and music in newborn infants', *Human Brain Mapping*, 31(4), pp. 595–603. Available at: https://doi.org/10.1002/hbm.20890.

Kotsoni, E., De Haan, M. and Johnson, M.H. (2001) 'Categorical Perception of Facial Expressions by 7-Month-Old Infants', *Perception*, 30(9), pp. 1115–1125. Available at: https://doi.org/10.1068/p3155.

Kuhl, P.K. (1981) 'Discrimination of speech by nonhuman animals: Basic auditory sensitivities conducive to the perception of speech-sound categories', *The Journal of the Acoustical Society of America*, 70(2), pp. 340–349. Available at: https://doi.org/10.1121/1.386782.

Kuhl, P.K. (2004) 'Early language acquisition: cracking the speech code', *Nature Reviews Neuroscience*, 5(11), pp. 831–843. Available at: https://doi.org/10.1038/nrn1533.

Kuhl, P.K. *et al.* (2014) ‘Infants’ brain responses to speech suggest Analysis by Synthesis’, *Proceedings of the National Academy of Sciences*, 111(31), pp. 11238–11245. Available at: https://doi.org/10.1073/pnas.1410963111.

Ladányi E, Persici V, Fiveash A, Tillmann B, Gordon RL. (2020). Is atypical rhythm a risk factor for develop-mental speech and language disorders? *WIREs Cogn. Sci.* 11(5):e1528

Lavan, N. *et al.* (2017) ‘Neural correlates of the affective properties of spontaneous and volitional laughter types’, *Neuropsychologia*, 95, pp. 30–39. Available at: https://doi.org/10.1016/j.neuropsychologia.2016.12.012.

Lavan, N., Scott, S.K. and McGettigan, C. (2016) ‘Laugh Like You Mean It: Authenticity Modulates Acoustic, Physiological and Perceptual Properties of Laughter’, *Journal of Nonverbal Behavior*, 40(2), pp. 133–149. Available at: https://doi.org/10.1007/s10919-015-0222-8.

Lecanuet, J-P., Granier-Deferre, C. (1993). Speech stimuli in 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

Lew, S. *et al.* (2013) ‘Effects of sutures and fontanels on MEG and EEG source analysis in a realistic infant head model’, *NeuroImage*, 76, pp. 282–293. Available at: https://doi.org/10.1016/j.neuroimage.2013.03.017.

Liberman, A.M. *et al.* (1967) ‘Perception of the speech code.’, *Psychological Review*, 74(6), pp. 431–461. Available at: https://doi.org/10.1037/h0020279.

Lillo-Martin, D.C. (1991). Universal Grammar andAmerican Sign Language: Setting the Null Argument Parameters (Boston: Kluwer Academic Publishers).

Lloyd-Fox, S. *et al.* (2014) ‘Coregistering functional near-infrared spectroscopy with underlying cortical areas in infants’, *Neurophotonics*, 1(2), p. 025006. Available at: https://doi.org/10.1117/1.NPh.1.2.025006.

Lloyd-Fox, S., Blasi, A. and Elwell, C.E. (2010) ‘Illuminating the developing brain: The past, present and future of functional near infrared spectroscopy’, *Neuroscience & Biobehavioral Reviews*, 34(3), pp. 269–284. Available at: https://doi.org/10.1016/j.neubiorev.2009.07.008.

Loukina, A. *et al.* (2011) ‘Rhythm measures and dimensions of durational variation in speech’, *The Journal of the Acoustical Society of America*, 129(5), pp. 3258–3270. Available at: https://doi.org/10.1121/1.3559709.

MacWhinney, B. (1992) ‘The CHILDES project: tools for analyzing talk’, *Child Language Teaching and Therapy*, 8(2), pp. 217–218. Available at: https://doi.org/10.1177/026565909200800211.

Mahmoudzadeh, M. *et al.* (2013) ‘Syllabic discrimination in premature human infants prior to complete formation of cortical layers’, *Proceedings of the National Academy of Sciences*, 110(12), pp. 4846–4851. Available at: https://doi.org/10.1073/pnas.1212220110.

Mampe, B. *et al.* (2009) ‘Newborns’ Cry Melody Is Shaped by Their Native Language’, *Current Biology*, 19(23), pp. 1994–1997. Available at: https://doi.org/10.1016/j.cub.2009.09.064.

Manfredi, C. *et al.* (2019) ‘Automated analysis of newborn cry: relationships between melodic shapes and native language’, *Biomedical Signal Processing and Control*, 53, p. 101561. Available at: https://doi.org/10.1016/j.bspc.2019.101561.

Maria, A. *et al.* (2018) ‘Emotional Processing in the First 2 Years of Life: A Review of Near-Infrared Spectroscopy Studies’, *Journal of Neuroimaging*, 28(5), pp. 441–454. Available at: https://doi.org/10.1111/jon.12529.

Marina Nespor *et al.* (2008) ‘Different Phrasal Prominence Realizations in VO and OV Languages’, *Lingue e linguaggio*, (2), pp. 139–168. Available at: https://doi.org/10.1418/28093.

Maris, E. and Oostenveld, R. (2007) ‘Nonparametric statistical testing of EEG- and MEG-data’, *Journal of Neuroscience Methods*, 164(1), pp. 177–190. Available at: https://doi.org/10.1016/j.jneumeth.2007.03.024.

Martinez-Alvarez, A. *et al.* (2023) ‘Newborns discriminate utterance-level prosodic contours’, *Developmental Science*, 26(2), p. e13304. Available at: https://doi.org/10.1111/desc.13304.

Mastropieri, D. and Turkewitz, G. (1999) ‘Prenatal experience and neonatal responsiveness to vocal expressions of emotion’, *Developmental Psychobiology*, 35(3), pp. 204–214. Available at: https://doi.org/10.1002/(SICI)1098-2302(199911)35:3<204::AID-DEV5>3.0.CO;2-V.

May, L. *et al.* (2011) ‘Language and the Newborn Brain: Does Prenatal Language Experience Shape the Neonate Neural Response to Speech?’, *Frontiers in Psychology*, 2. Available at: https://doi.org/10.3389/fpsyg.2011.00222.

May, L. *et al.* (2018) ‘The specificity of the neural response to speech at birth’, *Developmental Science*, 21(3), p. e12564. Available at: https://doi.org/10.1111/desc.12564.

Mazuka, R., Hasegawa, M. and Tsuji, S. (2014) ‘Development of non-native vowel discrimination: Improvement without exposure’, *Developmental Psychobiology*, 56(2), pp. 192–209. Available at: https://doi.org/10.1002/dev.21193.

Mazzocconi, C., Tian, Y. and Ginzburg, J. (2022) ‘What’s Your Laughter Doing There? A Taxonomy of the Pragmatic Functions of Laughter’, *IEEE Transactions on Affective Computing*, 13(3), pp. 1302–1321. Available at: https://doi.org/10.1109/TAFFC.2020.2994533.

McCauley, S.M. and Christiansen, M.H. (2019) ‘Language learning as language use: A cross-linguistic model of child language development.’, *Psychological Review*, 126(1), pp. 1–51. Available at: https://doi.org/10.1037/rev0000126.

McGettigan, C. *et al.* (2015) ‘Individual Differences in Laughter Perception Reveal Roles for Mentalizing and Sensorimotor Systems in the Evaluation of Emotional Authenticity’, *Cerebral Cortex*, 25(1), pp. 246–257. Available at: https://doi.org/10.1093/cercor/bht227.

Meek, J.H. *et al.* (1998) ‘Regional Hemodynamic Responses to Visual Stimulation in Awake Infants’, *Pediatric Research*, 43(6), pp. 840–843. Available at: https://doi.org/10.1203/00006450-199806000-00019.

Mehler, J. *et al.* (1988) ‘A precursor of language acquisition in young infants’, *Cognition*, 29(2), pp. 143–178. Available at: https://doi.org/10.1016/0010-0277(88)90035-2.

Mehler, J., Dupoux, E., Nazzi, T., Deheane-Lambertz, G. (1996). Coping with linguistic diversity : the infant’s viewpoint. In *Signal to Syntax : Bootstrapping from Speech to Grammar in Early Acquisition*, ed. JL Morgan, K Demuth, pp. 101-16. Mahwah, NJ : Lawrence Erlbaum Assoc.

Meyer, M. *et al.* (2005) ‘Distinct fMRI responses to laughter, speech, and sounds along the human peri-sylvian cortex’, *Cognitive Brain Research*, 24(2), pp. 291–306. Available at: https://doi.org/10.1016/j.cogbrainres.2005.02.008.

Minagawa-Kawai, Y. *et al.* (2011) ‘Optical Brain Imaging Reveals General Auditory and Language-Specific Processing in Early Infant Development’, *Cerebral Cortex*, 21(2), pp. 254–261. Available at: https://doi.org/10.1093/cercor/bhq082.

Mireault, G. *et al.* (2012) ‘Humour Perception and Creation between Parents and 3- to 6-month-old Infants’, *Infant and Child Development*, 21(4), pp. 338–347. Available at: https://doi.org/10.1002/icd.757.

Missana, M., Altvater-Mackensen, N. and Grossmann, T. (2017) ‘Neural correlates of infants’ sensitivity to vocal expressions of peers’, *Developmental Cognitive Neuroscience*, 26, pp. 39–44. Available at: https://doi.org/10.1016/j.dcn.2017.04.003.

Molfese, D.L., Freeman, R.B. and Palermo, D.S. (1975) ‘The ontogeny of brain lateralization for speech and nonspeech stimuli’, *Brain and Language*, 2, pp. 356–368. Available at: https://doi.org/10.1016/S0093-934X(75)80076-9.

Moon, C. (2017) ‘Prenatal Experience with the Maternal Voice’, in M. Filippa, P. Kuhn, and B. Westrup (eds) *Early Vocal Contact and Preterm Infant Brain Development*. Cham: Springer International Publishing, pp. 25–37. Available at: https://doi.org/10.1007/978-3-319-65077-7_2.

Moon, C., Cooper, R.P. and Fifer, W.P. (1993) ‘Two-day-olds prefer their native language’, *Infant Behavior and Development*, 16(4), pp. 495–500. Available at: https://doi.org/10.1016/0163-6383(93)80007-U.

Moon, C., Lagercrantz, H. and Kuhl, P.K. (2013) ‘Language experienced *in utero* affects vowel perception after birth: a two-country study’, *Acta Paediatrica*, 102(2), pp. 156–160. Available at: https://doi.org/10.1111/apa.12098.

Moore, B.C.J. (2008) ‘Basic auditory processes involved in the analysis of speech sounds’, *Philosophical Transactions of the Royal Society B: Biological Sciences*, 363(1493), pp. 947–963. Available at: https://doi.org/10.1098/rstb.2007.2152.

Morgan JL, Demuth K. (1996). *Signal to Syntax: Bootstrapping from Speech to Grammar in Early Acquisition*. Hillsdale, NJ: Lawrence Erlbaum

Most, T. and Michaelis, H. (2012) ‘Auditory, Visual, and Auditory–Visual Perceptions of Emotions by Young Children With Hearing Loss Versus Children With Normal Hearing’, *Journal of Speech, Language, and Hearing Research*, 55(4), pp. 1148–1162. Available at: https://doi.org/10.1044/1092-4388(2011/11-0060).

Nagy, E. and Molnar, P. (2004) ‘Homo imitans or homo provocans? Human imprinting model of neonatal imitation’, *Infant Behavior and Development*, 27(1), pp. 54–63. Available at: https://doi.org/10.1016/j.infbeh.2003.06.004.

Naik, A.N., Varadarajan, V.V. and Malhotra, P.S. (2021) ‘Early pediatric Cochlear implantation: An update’, *Laryngoscope Investigative Otolaryngology*, 6(3), pp. 512–521. Available at: https://doi.org/10.1002/lio2.574.

Nallet, C. and Gervain, J. (2021) ‘Neurodevelopmental Preparedness for Language in the Neonatal Brain’, *Annual Review of Developmental Psychology*, 3(1), pp. 41–58. Available at: https://doi.org/10.1146/annurev-devpsych-050620-025732.

Nallet, C. and Gervain, J. (2022) ‘Atypical neural oscillations in response to speech in infants and children with speech and language impairments: a systematic review’, *Hearing, Balance and Communication*, 20(3), pp. 145–154. Available at: https://doi.org/10.1080/21695717.2022.2084864.

Nallet, C., Berent, I., Werker, J. F., & Gervain, J. (2023). The neonate brain’s sensitivity to repetition-based structure: Specific to speech? *Developmental Science*, e13408. https://doi.org/10.1111/desc.13408

Narayan, C.R., Werker, J.F. and Beddor, P.S. (2010) ‘The interaction between acoustic salience and language experience in developmental speech perception: evidence from nasal place discrimination’, *Developmental Science*, 13(3), pp. 407–420. Available at: https://doi.org/10.1111/j.1467-7687.2009.00898.x.

Nazzi, T., Bertoncini, J. and Mehler, J. (1998) ‘Language discrimination by newborns: Toward an understanding of the role of rhythm.’, *Journal of Experimental Psychology: Human Perception and Performance*, 24(3), pp. 756–766. Available at: https://doi.org/10.1037/0096-1523.24.3.756.

Nelson, C.A. and De Haan, M. (1996) ‘Neural correlates of infants’ visual responsiveness to facial expressions of emotion’, *Developmental Psychobiology*, 29(7), pp. 577–

595. Available at: https://doi.org/10.1002/(SICI)1098-2302(199611)29:7<577::AID-DEV3>3.0.CO;2-R.

Nespor M, Vogel I. (1986). *Prosodic Phonology*, Vol. 28. Dordrecht, Neth.: Foris

Nespor M, Shukla M, van de Vijver R, Avesani C, Schraudolf H, Donati C. (2008). Different phrasal prominence realization in VO and OV languages. *Lingue Linguaggio* 7(2):139–68

Neumann, K. *et al.* (no date) 'A Survey on the Global Status of Newborn and Infant Hearing Screening'.

Obrig, H. (2010) 'From acoustic segmentation to language processing: evidence from optical imaging', *Frontiers in Neuroenergetics* [Preprint]. Available at: https://doi.org/10.3389/fnene.2010.00013.

Ojemann, G. *et al.* (1989) 'Cortical language localization in left, dominant hemisphere: An electrical stimulation mapping investigation in 117 patients', *Journal of Neurosurgery*, 71(3), pp. 316–326. Available at: https://doi.org/10.3171/jns.1989.71.3.0316.

Olulade, O.A. *et al.* (2020) 'The neural basis of language development: Changes in lateralization over age', *Proceedings of the National Academy of Sciences*, 117(38), pp. 23477–23483. Available at: https://doi.org/10.1073/pnas.1905590117.

Ortiz Barajas, M.C., Guevara, R. and Gervain, J. (2021) 'The origins and development of speech envelope tracking during the first months of life', *Developmental Cognitive Neuroscience*, 48, p. 100915. Available at: https://doi.org/10.1016/j.dcn.2021.100915.

Ortiz-Barajas, M.C., Guevara, R. and Gervain, J. (2023) 'Neural oscillations and speech processing at birth', *iScience*, 26(11), p. 108187. Available at: https://doi.org/10.1016/j.isci.2023.108187.

Otsuka, Y. *et al.* (2007) 'Neural activation to upright and inverted faces in infants measured by near infrared spectroscopy', *NeuroImage*, 34(1), pp. 399–406. Available at: https://doi.org/10.1016/j.neuroimage.2006.08.013.

Panneton, R.K., De Casper, A.J. (1986). *Newborns' postnatal preference for a prenatally experienced melody*. Paper presented at the International Conference on Infant Studies, Beverly Hills, CA

Papoušek, M. *et al.* (1990) 'Infant responses to prototypical melodic contours in parental speech', *Infant Behavior and Development*, 13(4), pp. 539–545. Available at: https://doi.org/10.1016/0163-6383(90)90022-Z.

Peña, M. *et al.* (2003) 'Sounds and silence: An optical topography study of language recognition at birth', *Proceedings of the National Academy of Sciences*, 100(20), pp. 11702–11705. Available at: https://doi.org/10.1073/pnas.1934290100.

Peña, M., Pittaluga, E. and Mehler, J. (2010) 'Language acquisition in premature and full-term infants', *Proceedings of the National Academy of Sciences*, 107(8), pp. 3823–3828. Available at: https://doi.org/10.1073/pnas.0914326107.

Peña, M., Werker, J.F. and Dehaene-Lambertz, G. (2012) 'Earlier Speech Exposure Does Not Accelerate Speech Acquisition', *The Journal of Neuroscience*, 32(33), pp. 11159–11163. Available at: https://doi.org/10.1523/JNEUROSCI.6516-11.2012.

Petitjean, C. (1989). *Une condition de l'audition fœtale : la conduction sonore osseuse. Conséquences cliniques et applications pratiques*. MD Miss., Univ. Franche-Comté, Besançon, Fr.

Pierce LJ, Klein D, Chen J-K, Delcenserie A, Genesee F. (2014). Mapping the unconscious maintenance of a lost first language. *PNAS* 111(48):17314–19

Pimperton, H. *et al.* (2016) 'The impact of universal newborn hearing screening on long-term literacy outcomes: a prospective cohort study', *Archives of Disease in Childhood*, 101(1), pp. 9–15. Available at: https://doi.org/10.1136/archdischild-2014-307516.

Polka, L., Masapollo, M. and Ménard, L. (2022) 'Setting the Stage for Speech Production: Infants Prefer Listening to Speech Sounds With Infant Vocal Resonances', *Journal of Speech, Language, and Hearing Research*, 65(1), pp. 109–120. Available at: https://doi.org/10.1044/2021_JSLHR-21-00412.

Prince, P. *et al.* (2021) 'Neural correlates of visual stimulus encoding and verbal working memory differ between cochlear implant users and normal-hearing controls', *European Journal of Neuroscience*, 54(3), pp. 5016–5037. Available at: https://doi.org/10.1111/ejn.15365.

Provine, R.R. (2004) 'Laughing, Tickling, and the Evolution of Speech and Self', *Current Directions in Psychological Science*, 13(6), pp. 215–218. Available at: https://doi.org/10.1111/j.0963-7214.2004.00311.x.

Purcell, P.L. *et al.* (2021) 'Cochlear Implantation in Infants: Why and How', *Trends in Hearing*, 25, p. 233121652110317. Available at: https://doi.org/10.1177/23312165211031751.

Querleu, D. *et al.* (1988) 'Fetal hearing', *European Journal of Obstetrics & Gynecology and Reproductive Biology*, 28(3), pp. 191–212. Available at: https://doi.org/10.1016/0028-2243(88)90030-5.

Ramus, F. *et al.* (2000) 'Language Discrimination by Human Newborns and by Cotton-Top Tamarin Monkeys', *Science*, 288(5464), pp. 349–351. Available at: https://doi.org/10.1126/science.288.5464.349.

Ramus, F., Nespor, M. and Mehler, J. (2000) 'Correlates of linguistic rhythm in the speech signal', *Cognition*, 75(1), pp. AD3–AD30. Available at: https://doi.org/10.1016/S0010-0277(00)00101-3.

Roche-Labarbe, N. *et al.* (2008) 'High-resolution electroencephalography and source localization in neonates', *Human Brain Mapping*, 29(2), pp. 167–176. Available at: https://doi.org/10.1002/hbm.20376.

Saffran, J.R. *et al.* (2007) 'Dog is a dog is a dog: Infant rule learning is not specific to language', *Cognition*, 105(3), pp. 669–680. Available at: https://doi.org/10.1016/j.cognition.2006.11.004.

Sansavini, A., Bertoncini, J. and Giovanelli, G. (1997) 'Newborns discriminate the rhythm of multisyllabic stressed words.', *Developmental Psychology*, 33(1), pp. 3–11. Available at: https://doi.org/10.1037/0012-1649.33.1.3.

Sato, H. *et al.* (2012) 'Cerebral hemodynamics in newborn infants exposed to speech sounds: A whole-head optical topography study', *Human Brain Mapping*, 33(9), pp. 2092–2103. Available at: https://doi.org/10.1002/hbm.21350.

Saurini, P., Nola, G. and Lendvai, D. (no date) 'Otoacoustic emissions: a new method for newborn hearing screening'.

Scott, S.K. *et al.* (2014) 'The social life of laughter', *Trends in Cognitive Sciences*, 18(12), pp. 618–620. Available at: https://doi.org/10.1016/j.tics.2014.09.002.

Segal, O., Houston, D. and Kishon-Rabin, L. (2016) 'Auditory Discrimination of Lexical Stress Patterns in Hearing-Impaired Infants with Cochlear Implants Compared with Normal Hearing: Influence of Acoustic Cues and Listening Experience to the Ambient Language', *Ear & Hearing*, 37(2), pp. 225–234. Available at: https://doi.org/10.1097/AUD.0000000000000243.

Selkirk E. 1986. On derived domains in sentence phonology. *Phonology* 3:371–405

Sharma, S.D. *et al.* (2020) 'Hearing and speech benefits of cochlear implantation in children: A review of the literature', *International Journal of Pediatric Otorhinolaryngology*, 133, p. 109984. Available at: https://doi.org/10.1016/j.ijporl.2020.109984.

Shi, F. *et al.* (2011) 'Infant Brain Atlases from Neonates to 1- and 2-Year-Olds', *PLoS ONE*. Edited by H. Okazawa, 6(4), p. e18746. Available at: https://doi.org/10.1371/journal.pone.0018746.

Shukla M, White KS, Aslin RN. (2011). Prosody guides the rapid mapping of auditory word forms onto visual objects in 6-mo-old infants. *PNAS* 108(15):6038–43

Simner, M.L. (1971) 'Newborn's response to the cry of another infant.', *Developmental Psychology*, 5(1), pp. 136–150. Available at: https://doi.org/10.1037/h0031066.

Singh, L., Morgan, J.L. and Best, C.T. (2002) 'Infants' Listening Preferences: Baby Talk or Happy Talk?', *Infancy*, 3(3), pp. 365–394. Available at: https://doi.org/10.1207/S15327078IN0303_5.

Slocombe, K.E. *et al.* (2022) 'Chimpanzee vocal communication: what we know from the wild', *Current Opinion in Behavioral Sciences*, 46, p. 101171. Available at: https://doi.org/10.1016/j.cobeha.2022.101171.

Spence, M.J. and DeCasper, A.J. (1987) 'Prenatal experience with low-frequency maternal-voice sounds influence neonatal perception of maternal voice samples', *Infant Behavior and Development*, 10(2), pp. 133–142. Available at: https://doi.org/10.1016/0163-6383(87)90028-2.

Sroufe, L.A. and Wunsch, J.P. (1972) 'The Development of Laughter in the First Year of Life', *Child Development*, 43(4), p. 1326. Available at: https://doi.org/10.2307/1127519.

Stetson, R.H., (1928), Motor phonetics: A study of speech movements in action. *Archives Néerlandaises de Phonétique Expérimentale 3*, (Réedition 1987 ; J.A.S. KELSO K.G. MUNHALL, Boston: Little, Brown and Company), 1 - 216.

Swingley, D. (2022) 'Infants' Learning of Speech Sounds and Word-Forms', in Swingley, D., *The Oxford Handbook of the Mental Lexicon*. Edited by A. Papafragou, J. C. Trueswell, and L. R. Gleitman. Oxford University Press, pp. 266–291. Available at: https://doi.org/10.1093/oxfordhb/9780198845003.013.6.

Szameitat, D.P. *et al.* (2010) 'It is not always tickling: Distinct cerebral responses during perception of different laughter types', *NeuroImage*, 53(4), pp. 1264–1271. Available at: https://doi.org/10.1016/j.neuroimage.2010.06.028.

Taylor, C.L. *et al.* (2013) 'Risk Factors for Children's Receptive Vocabulary Development from Four to Eight Years in the Longitudinal Study of Australian Children', *PLoS ONE*. Edited by K. Paterson, 8(9), p. e73046. Available at: https://doi.org/10.1371/journal.pone.0073046.

Telkemeyer, S. *et al.* (2009) 'Sensitivity of Newborn Auditory Cortex to the Temporal Structure of Sounds', *The Journal of Neuroscience*, 29(47), pp. 14726–14733. Available at: https://doi.org/10.1523/JNEUROSCI.1246-09.2009.

Tomasello, M. (2000) 'Do young children have adult syntactic competence?', *Cognition*, 74(3), pp. 209–253. Available at: https://doi.org/10.1016/S0010-0277(99)00069-4.

Valenza, E. *et al.* (1996) 'Face preference at birth.', *Journal of Experimental Psychology: Human Perception and Performance*, 22(4), pp. 892–903. Available at: https://doi.org/10.1037/0096-1523.22.4.892.

Van Straaten, H. (1999) 'Automated auditory brainstem response in neonatal hearing screening', *Acta Paediatrica*, 88(s432), pp. 76–79. Available at: https://doi.org/10.1111/j.1651-2227.1999.tb01165.x.

Vannasing, P. *et al.* (2016) 'Distinct hemispheric specializations for native and non-native languages in one-day-old newborns identified by fNIRS', *Neuropsychologia*, 84, pp. 63–69. Available at: https://doi.org/10.1016/j.neuropsychologia.2016.01.038.

Ventureyra VAG, Pallier C, Yoo H-Y. (2004). The loss of first language phonetic perception in adopted Koreans. *J. Neurolinguist.* 17(1):79–91

Vettin, J. and Todt, D. (2004) ‘Laughter in Conversation: Features of Occurrence and Acoustic Structure’, *Journal of Nonverbal Behavior*, 28(2), pp. 93–115. Available at: https://doi.org/10.1023/B:JONB.0000023654.73558.72.

Viinikainen, M., Kätsyri, J. and Sams, M. (2012) ‘Representation of perceived sound valence in the human brain’, *Human Brain Mapping*, 33(10), pp. 2295–2305. Available at: https://doi.org/10.1002/hbm.21362.

Viola Macchi, C., Turati, C. and Simion, F. (2004) ‘Can a Nonspecific Bias Toward Top-Heavy Patterns Explain Newborns’ Face Preference?’, *Psychological Science*, 15(6), pp. 379–383. Available at: https://doi.org/10.1111/j.0956-7976.2004.00688.x.

Vohr, B. (2014) ‘Speech and language outcomes of very preterm infants’, *Seminars in Fetal and Neonatal Medicine*, 19(2), pp. 78–83. Available at: https://doi.org/10.1016/j.siny.2013.10.007.

Vouloumanos, A. *et al.* (2010) ‘The Tuning of Human Neonates’ Preference for Speech’, *Child Development*, 81(2), pp. 517–527. Available at: https://doi.org/10.1111/j.1467-8624.2009.01412.x.

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

Wallois, F. *et al.* (2012) ‘Usefulness of simultaneous EEG–NIRS recording in language studies’, *Brain and Language*, 121(2), pp. 110–123. Available at: https://doi.org/10.1016/j.bandl.2011.03.010.

Warren, J.E. *et al.* (2006) ‘Positive Emotions Preferentially Engage an Auditory–Motor “Mirror” System’, *The Journal of Neuroscience*, 26(50), pp. 13067–13075. Available at: https://doi.org/10.1523/JNEUROSCI.3907-06.2006.

Weikum, W.M. *et al.* (2012) ‘Prenatal exposure to antidepressants and depressed maternal mood alter trajectory of infant speech perception’, *Proceedings of the National Academy of Sciences*, 109(supplement_2), pp. 17221–17227. Available at: https://doi.org/10.1073/pnas.1121263109.

Werker, J.F. (2018) ‘Perceptual beginnings to language acquisition’, *Applied Psycholinguistics*, 39(4), pp. 703–728. Available at: https://doi.org/10.1017/S0142716418000152.

Werker, J.F. and Curtin, S. (2005) ‘PRIMIR: A Developmental Framework of Infant Speech Processing’, *Language Learning and Development*, 1(2), pp. 197–234. Available at: https://doi.org/10.1080/15475441.2005.9684216.

Werker, J.F. and Tees, R.C. (1984) 'Cross-language speech perception: Evidence for perceptual reorganization during the first year of life', *Infant Behavior and Development*, 7(1), pp. 49–63. Available at: https://doi.org/10.1016/S0163-6383(84)80022-3.

Werker, J.F. and Tees, R.C. (2005) 'Speech perception as a window for understanding plasticity and commitment in language systems of the brain', *Developmental Psychobiology*, 46(3), pp. 233–251. Available at: https://doi.org/10.1002/dev.20060.

Wermke, K. and Mende, W. (2009) 'Musical elements in human infants' cries: In the beginning is the melody', *Musicae Scientiae*, 13(2_suppl), pp. 151–175. Available at: https://doi.org/10.1177/1029864909013002081.

Wernicke, C. (1874). *Der Aphasische Symptomencomplex—Eine Psychologische Studie auf Anatomischer Basis*, (Cohn & Weigert, Breslau, Germany).

Wiethoff, S. *et al.* (2008) 'Cerebral processing of emotional prosody—influence of acoustic parameters and arousal', *NeuroImage*, 39(2), pp. 885–893. Available at: https://doi.org/10.1016/j.neuroimage.2007.09.028.

Wiget, L. *et al.* (2010) 'How stable are acoustic metrics of contrastive speech rhythm?', *The Journal of the Acoustical Society of America*, 127(3), pp. 1559–1569. Available at: https://doi.org/10.1121/1.3293004.

Wildgruber, D. *et al.* (2002) 'Dynamic Brain Activation during Processing of Emotional Intonation: Influence of Acoustic Parameters, Emotional Valence, and Sex', *NeuroImage*, 15(4), pp. 856–869. Available at: https://doi.org/10.1006/nimg.2001.0998.

Wolff, P. H. (1963). *Observations on the early development of smiling*. In B. M. Foss (Ed.). *Determinants of Infant Behavior*, Vol. 2. New York: Wiley.

Wolke, D. *et al.* (2008) 'Specific Language Difficulties and School Achievement in Children Born at 25 Weeks of Gestation or Less', *The Journal of Pediatrics*, 152(2), pp. 256-262.e1. Available at: https://doi.org/10.1016/j.jpeds.2007.06.043.

Yeung, H.H. and Werker, J.F. (2013) 'Lip Movements Affect Infants' Audiovisual Speech Perception', *Psychological Science*, 24(5), pp. 603–612. Available at: https://doi.org/10.1177/0956797612458802.

Yoshinaga-Itano, C. *et al.* (2008) 'Outcomes of Children with Mild Bilateral Hearing Loss and Unilateral Hearing Loss', *Seminars in Hearing*, 29(2), pp. 196–211. Available at: https://doi.org/10.1055/s-2008-1075826.

Zatorre, R.J. (2001) 'Spectral and Temporal Processing in Human Auditory Cortex', *Cerebral Cortex*, 11(10), pp. 946–953. Available at: https://doi.org/10.1093/cercor/11.10.946.

Zhang, D. *et al.* (2017) 'Discrimination of emotional prosodies in human neonates: A pilot fNIRS study', *Neuroscience Letters*, 658, pp. 62–66. Available at: https://doi.org/10.1016/j.neulet.2017.08.047.

Appendix A

Affirmative
Lili always has breakfast before going to work.
Ten-year-old Much Middleton was helping his father.
I wonder what kind of cookies she's going to make.
Jonathan had to take the sheep to the grassy fields.
Interrogative
Should we put the little baby in the blue highchair?
Do you want to go to the library on Monday?
Do you think that your daddy can sew the teddy bear?
Do you want a white paper towel to wipe your hands?
Have you ever been to the zoo and seen the zebra?
Do you remember the walruses on mummys' shirt?
Exclamative
We shall go to the seaside tomorrow afternoon!
What extraordinary courage this lady showed!
Look at this beautiful bird flying in the blue sky!
Our summer vacation will start in just three weeks!
Why don't you look at some of the toys in the basket!
How confident and beautiful that gentleman is!

Table A.1 : English sentences from experiment 4.

Affirmative
Questa medaglietta è gialla e rossa.
Guglielmo ama molto le piccolo volpi.
Il nonno prepara una torta di mele.
Gli orsi dormono tutti in una casa.
Interrogative
Questa pallina qui chi te l'ha regalata ?
Lo sai come si chiama quell' animale lì ?
Dove l'abbiamo trovata la lucertola ?
Come fai a montare sulla grande scala ?
Che avete fatto a casa di Monica ?
Che cosa gli piace fare la domenica?
Exclamative
Un uccellino bello con le ali rosse!
Prendi quello tuo che un po' più grosso !
Un coniglietto carino corre nel prato !
Avevi avuto una bella idea !
Guarda come saltano i piccolo cani!
Allora prendine un altro che sia buono!

Table A.2 : Italian sentences from experiment 4.