UNIVERSITÉ DE PARIS

Ecole doctorale 261: Cognition, Comportements, Conduites Humaines

Integrative Neuroscience and Cognition Center

Oscillations neuronales et traitement de la parole pendant les premiers mois de vie

---

Neural oscillations and speech processing during the first months of life

Thèse de Doctorat en Neurosciences Cognitives

Maria Clemencia Ortiz Barajas

Date de soutenance: 27 novembre 2019

Membres de Jury:

Judit Gervain	Directrice de Recherche CNRS	Directrice de thèse
Christian Lorenzi	Professeur des Universités	Rapporteur
Claudia Männel	Research Group Leader	Rapporteur
Marcela Peña Garay	Professeur des Universités	Examinateur
Laura Dugué	Maître de Conférences (HDR)	Examinateur

To Gabriela, you put the rhythm in my life

Résumé

Les oscillations neuronales correspondent à l'activité synchronisée d'ensembles neuronaux, qui peuvent être couplés spontanément ou par un stimulus commun. Certaines théories proposent que les oscillations neuronales à certaines bandes de fréquence seraient en lien avec la perception de la parole et le traitement du langage. Cette hypothèse est fondée sur la correspondance entre les échelles temporelles des unités langagières et les fréquences des oscillations. Les oscillations de la bande gamma basse (25-35 Hz), thêta (4-8 Hz) et delta (1-3 Hz) qui seraient impliquées dans le traitement des unités sub(phonémiques), syllabiques et prosodiques, respectivement. Actuellement, nous ignorons si ces oscillations sont déjà présentes à la naissance, et si elles sont en lien avec le traitement du langage dès le début du développement. L'objectif de cette thèse est d'étudier les oscillations neuronales et leur relation avec le traitement du langage dans les premiers mois de la vie. A cette fin, nous avons examiné le fonctionnement des oscillations neuronales chez le nouveau-né, qui ont une expérience langagière (extra-uterine) limitée. Nous avons également examiné ce fonctionnement à l'âge de 6 mois pour tester l'impact de l'expérience langagière (postnatale) sur les oscillations.

Cette thèse est organisée en trois parties. La première partie est dédiée à l'enregistrement en électroencéphalographie (EEG) de l'activité neuronal des nouveau-nés en état de repos (*resting state*) aussi bien que pendant l'écoute de phrases dans la langue maternelle (français), dans une langue inconnue, mais rythmiquement similaire (espagnol) et dans une langue inconnue à un rythme différent (anglais). Nous avons réalisé des analyses en temps-fréquence afin d'étudier les oscillations en état de repos et pendant la stimulation langagière. Les résultats montrent une activation dans les bandes thêta et gamma basse pendant le traitement de la parole, et cette activation est modulée par la familiarité avec la langue présentée. En outre, nous avons examiné la synchronisation des oscillations neuronales avec l'enveloppe temporelle de la parole, un phénomène connu comme *envelope tracking*. En effet, cette capacité serait nécessaire pour la compréhension. Analyser la performance des nouveau-nés, dont les connaissances linguistiques sont limitées, permet de mieux comprendre le rôle causal de l'envelope tracking dans la compréhension. Nous avons trouvé que les nouveau-nés sont capables d'encoder l'enveloppe temporelle de la parole autant dans la langue maternelle que dans les deux langues inconnues.

La deuxième partie de cette thèse s'intéresse au développement des oscillations neuronales pendant les premiers mois de la vie. Nous avons enregistré l'EEG chez le

nourrisson de 6 mois dans la même tâche d'écoute (français, espagnol, anglais) qu'à la naissance. L'analyse des oscillations neuronales a montré l'émergence des oscillations gamma haute ainsi que l'effet de la familiarité de la langue dans les bandes thêta et gamma (hautes et basses). De plus, les résultats de l'enveloppe tracking montrent que l'encodage de l'enveloppe temporelle à cet âge est différent entre les trois langues, l'activité neuronale montrant la meilleure corrélation avec l'enveloppe de l'anglais. Nous faisons l'hypothèse que cette différence est due aux caractéristiques acoustiques des trois langues et non à la familiarité avec la langue maternelle.

Dans la troisième partie, nous examinons si les oscillations neuronales peuvent servir de prédicteurs précoces du développement langagier. A cette fin, nous avons suivi de façon longitudinale les enfants testés dans l'étude EEG à la naissance et à 6 mois et nous avons évalué leur vocabulaire à 12 et à 18 mois. Nous n'avons pas trouvé de corrélations fortes entre les mesures neuronales et la taille du vocabulaire. Cependant, certains effets potentiels ont émergé et méritent d'être confirmés dans de futures études.

En conclusion, ce travail de recherche contribue à une meilleure compréhension du rôle des oscillations neuronales dans le traitement de la parole et l'acquisition du langage. Nous avons démontré comment les oscillations neuronales changent pendant les 6 premiers mois de la vie, facilitant l'acquisition précoce du langage.

Mots-clés: oscillations neuronales, envelope tracking, EEG, nouveau-nés, nourrissons

Abstract

Neural oscillations correspond to the synchronous activity generated by neuronal assemblies, which can be naturally coupled, or coupled by a common stimulus. It has been proposed that neural oscillations at certain frequency bands are involved in speech processing. This is partly due to the close correspondence between the time scales of speech units and the frequencies at which these oscillations occur. More specifically low-gamma (25-35 Hz), theta (4-8 Hz) and delta (1-3 Hz) oscillations have been argued to aid in the processing of (sub)phonemic, syllabic and phrasal units in the speech signal, respectively. Currently it is unknown whether neural oscillations are already in place at birth, and if they are, whether they relate to speech processing. The aim of this thesis is to investigate neural oscillations and their relationship with speech processing at the very beginning of life. To do so, we assessed how neural oscillations operate in the newborn brain, when experience with the (extrauterine) speech signal is limited, and in the brain of 6-month-olds, once infants have gained (postnatal) experience with spoken language.

This work is organized in three main parts. In the first part, we recorded the neural activity of newborns using electroencephalography (EEG) during rest and when listening to sentences in a familiar language (heard prenatally), as well as in a rhythmically similar and in a rhythmically different unfamiliar language. We implemented time-frequency analysis to investigate the effect of speech stimulation on resting state neural oscillations, as well as the effect of language familiarity. Additionally, we also investigated the synchronization of neural oscillations to the speech envelope, a phenomenon known as speech envelope tracking. It has been argued that speech envelope tracking is necessary for comprehension to take place. By assessing whether newborns, a population with sophisticated speech processing abilities but very little knowledge of language, perform speech envelope tracking, we could shed light on its relationship with speech comprehension. Results show that oscillations in the theta and low-gamma bands seem to be modulated during speech processing, and that this modulation depends on language familiarity. Additionally, newborns prenatally exposed only to French possess the neural ability to encode the speech envelope in familiar and unfamiliar languages equally well.

In the second part of this study, we investigated how neural oscillations develop in the first months of life. We recorded the neural activity of 6-month-olds (using EEG) while they listened to sentences in their native language, as well as in the same unfamiliar languages as at birth. When assessing neural oscillations we found the emergence of high-gamma

oscillations, and the effect of language familiarity on the theta and gamma (low and high) bands. When assessing speech envelope tracking we found that encoding of the speech envelope at this age differs slightly across the tested languages, but this seems to be related to acoustics and not to language familiarity.

For the third part, we investigated the potential of neural oscillations as predictors of language outcome. To do so, we followed participants longitudinally, and after the EEG recordings at birth and at 6 months, we collected information about their vocabulary size at 12 and 18 months. When assessing whether the early neural measures could predict the later language outcomes, we found no evidence for a strong predictive relationship. However some suggestive effects were identified that warrant further exploration and confirmation in future studies.

Overall, this work contributes to a better understanding of the role of neural oscillations in speech processing and language acquisition. We show how neural oscillations change after the first 6 months of life, subserving early language discrimination and the beginning of the attunement to the native language.

Keywords: neural oscillations, speech envelope tracking, EEG, newborns, infants

The present thesis has led to the following publications:

Ortiz-Barajas, M. C., Guevara Erra, R., Gervain, J. (under revision). Newborns can encode the speech envelope in familiar and unfamiliar languages.

Ortiz-Barajas, M. C., Gervain, J. (accepted). The role of prenatal experience and basic auditory mechanisms in the development of language. Accepted with minor revisions in *Minnesota Symposia on Child Psychology*, Volume 40: Human Communication: Origins, Mechanisms and Functions, Editors: M. Sera and M. Koenig. Wiley E-book, Hoboken, NJ. The projected publication year is 2020.

Varnet, L., Ortiz-Barajas, M. C., Erra, R. G., Gervain, J., and Lorenzi, C. (2017). A cross-linguistic study of speech modulation spectra. *J. Acoust. Soc. Am.* **142**(4), 1976–1989.

Acknowledgements

This thesis is a token of love, trust and patience, because it has been thanks to the love, trust and patience from many people that this thesis can finally see the light.

First, I would like to thank the families. If it were not for their trust and selflessness, this work would have been impossible to achieve. Thank you for allowing your babies to participate, and for having the motivation and the patience to accompany us for 18 months.

Thank you Judit for having taking me on board on this project, it has been very enriching having you as my mentor. Thank you for transmitting me your love for research and knowledge, you have taught how truly fascinating the infant brain is. Thank you for guiding me but also for trusting me to do things my own way. Thank you for always being positive and supportive, and always valorising my work. You have supported me professionally, academically and personally, and I am very grateful for that. Gracias también a ti Ramón por las buenas conversaciones, por las discusiones metodológicas y por tu ayuda con los análisis.

I would like to take this opportunity to thank the members of my jury, Christian Lorenzi, Claudia Männel, Marcela Peña, and Laura Dugué, for their careful reading of my manuscript, and for their feedback and suggestions. I am really honoured that you have accepted to be my jury.

This thesis was funded by the Predictable network. I am very grateful for their financial funding, for all the training opportunities, and for making me part of an incredible research network.

I am also grateful to have great colleagues and an amazing work environment. Thank you for the team meetings that provide us with a knowledgeable atmosphere where to discuss our work. Thank you to the members of the Babylab, especially to Viviane, Maxine and Anthony, for taking care of the Babylab so beautifully, and for always been available and happy to help.

I would also like to thank the persons who helped me with the recruitment. Thank you Lucie for not only helping me recruit at the maternity, but also for being amazing company during those long testing days. Thanks to the crash French course that you gave me I was able

to recruit by myself one day. I also would like to thank Anouche and Olivier, who helped me with the recruitment and the logistics of my longitudinal study.

I am truly grateful for the girls from the *bureau de doctorants*. Thank you, Elena, Oytun, Mélanie, Caterina, and Andrea, for supporting me from the first day I arrive at the office. For helping me navigate my life in France and for making our office a happy and friendly environment.

Gracias a mis padres y a mi hermana, que aunque físicamente lejos, siento cada día cerca. Gracias por su amor incondicional, por su apoyo constante, y por haberme dado tantas oportunidades en la vida, que me han formado para conseguir lo que quiero.

Thank you Jonathan, you have been the most amazing support in this adventure. From starting a new life in France so I could start this project, to being my rock day by day so I could complete it. Your love has been fundamental to this accomplishment. Gabrielita, gracias también a ti hija por tu paciencia, tu amor, y tus besos y abrazos reconfortantes que me recargan de energía cada día.

Table of Contents

CHAPTER 1: GENERAL INTRODUCTION 1

1.1 NEURAL OSCILLATIONS 4

1.1.1 NEURAL OSCILLATIONS IN ANIMALS 5

1.1.2 NEURAL OSCILLATIONS AT REST 5

1.1.3 NEURAL OSCILLATIONS AND SPEECH PROCESSING 6

1.1.4 NEURAL OSCILLATIONS AND LANGUAGE COMPREHENSION 8

1.2 SPEECH ENVELOPE TRACKING 9

1.2.1 DEFINITION OF SPEECH ENVELOPE 10

1.2.2 SPEECH ENVELOPE AND COMPREHENSION 11

1.2.3 MEASURES OF ENVELOPE TRACKING 12

1.3 SPEECH PERCEPTION AND LANGUAGE PROCESSING DURING INFANCY 14

1.3.1 PRENATAL HEARING 14

1.3.2 SPEECH PERCEPTION AT BIRTH 15

1.3.3 SPEECH PERCEPTION DURING THE FIRST YEAR OF LIFE 16

1.4 PREDICTING LANGUAGE OUTCOME 20

1.5 RESEARCH QUESTIONS AND HYPOTHESES 22

CHAPTER 2: NEURAL OSCILLATIONS IN THE HUMAN NEWBORN BRAIN 25

2.1 INTRODUCTION 27

2.2 MATERIALS AND METHODS 29

2.2.1 PARTICIPANTS 29

2.2.2 PROCEDURE 30

2.2.3 STIMULI 30

2.2.4 EEG DATA ACQUISITION 32

2.2.5 ENVELOPE TRACKING ANALYSIS 32

2.2.6 TIME-FREQUENCY ANALYSIS 37

2.3 RESULTS 40

2.3.1 ENVELOPE TRACKING 40

2.3.2 TIME-FREQUENCY ANALYSIS 46

2.4 DISCUSSION 53

CHAPTER 3: NEURAL OSCILLATIONS IN SIX-MONTH-OLD INFANTS 59

3.1 INTRODUCTION 61

3.2 MATERIALS AND METHODS 62

3.2.1 PARTICIPANTS 62

3.2.2 PROCEDURE 63

3.2.3 STIMULI 64

3.2.4 EEG DATA ACQUISITION 64

3.2.5 ENVELOPE TRACKING ANALYSIS 64

3.2.6 TIME-FREQUENCY ANALYSIS 67

3.3 RESULTS 69

3.3.1 ENVELOPE TRACKING 69

3.3.2 TIME-FREQUENCY ANALYSIS 79

3.4 DISCUSSION 85

CHAPTER 4: PREDICTING LANGUAGE OUTCOME 95

4.1 INTRODUCTION 97

4.2 MATERIALS AND METHODS 98

4.2.1 PARTICIPANTS 98

4.2.2 PROCEDURE 98

4.2.3 ANALYSIS OF LONGITUDINAL DATA 100

4.3 RESULTS 101

4.3.1 VOCABULARY GROWTH 101

4.3.2 PREDICTING LANGUAGE DEVELOPMENT 101

4.4 DISCUSSION 104

CHAPTER 5: GENERAL DISCUSSION 107

5.1 SUMMARY OF RESULTS 109

5.1.1 RESTING STATE NEURAL OSCILLATIONS 109

5.1.2 NEURAL OSCILLATIONS AND SPEECH PROCESSING 109

5.1.3 SPEECH ENVELOPE TRACKING 110

5.1.4 LINKAGE BETWEEN NEURAL OSCILLATIONS AND LANGUAGE OUTCOME 111

5.2 IMPLICATIONS FOR LANGUAGE LEARNING 111

5.2.1 RESTING STATE NEURAL OSCILLATIONS 111

5.2.2 NEURAL OSCILLATIONS AND SPEECH PROCESSING 112

5.2.3 DEVELOPMENTAL CHANGES IN NEURAL OSCILLATIONS 114

5.2.4 SPEECH ENVELOPE TRACKING 115
5.2.5 LINKAGE BETWEEN NEURAL OSCILLATIONS AND LANGUAGE OUTCOME 117
5.3 LIMITATIONS 117
5.4 FUTURE DIRECTIONS 118
5.5 GENERAL CONCLUSION 119

REFERENCES 121

CHAPTER 1: General Introduction

Most children acquire their native language(s) rapidly and effortlessly during the first years of life regardless of culture. They go from *cooing* and *gooing* at 3 months, to uttering their first words by the end of the first year. At the age of 3 years they speak in full sentences and can be understood by most adults. But how do they learn so easily? This puzzling question has given rise to decades of behavioural research that has allowed us to gain important insights into how these abilities are mastered early in childhood.

At the beginning of life infants are equipped with a functional (albeit not fully mature) auditory system, and a set of perceptual abilities that help them learn any language that they hear in their environment. However, language acquisition starts before birth, and this manifests itself in some of the abilities they exhibit shortly after birth: newborns prefer their mother's voice over other female voices (DeCasper & Fifer, 1980), as well as the language their mother spoke during pregnancy over other, rhythmically different languages (Mehler et al., 1988; Moon et al., 1993). They also recognize a story heard frequently during the last weeks of gestation (DeCasper & Spence, 1986).

Up to date, relatively little is known about how these sophisticated speech processing abilities operate in the developing human brain and about the neural architecture that supports them. This knowledge gap motivated our thesis/doctoral project: we sought to explore the neural mechanisms underlying speech processing in the infant brain.

We have chosen to investigate one specific type of neural mechanism, neural oscillations, which have gained a lot of attention in cognitive neuroscience over the last two decades. Neural oscillations correspond to the synchronous activity of neuronal assemblies, which can be spontaneously coupled, or coupled by a common stimulus (Arnal et al., 2015). Neural oscillations have been proposed to be involved in speech processing, among many other cognitive functions (Giraud & Poeppel, 2012). More specifically low-gamma (25-35 Hz), theta (4-8 Hz) and delta oscillations (1-3 Hz) have been argued to aid in the processing of phonemic, syllabic and phrasal units in the speech signal, respectively (Giraud & Poeppel, 2012).

The general objective of this dissertation is to investigate neural oscillations and their relationship with speech processing at the very beginning of life. Currently it is unknown whether neural oscillations are already in place at birth, or whether they require experience with the (extrauterine) speech signal to emerge. To answer this question, we assessed how neural oscillations operate in the newborn brain during rest and when listening to sentences in a familiar language (heard prenatally), as well as in unfamiliar languages (Chapter 2). These

different testing conditions allowed us to evaluate the endogenous occurring oscillations that are in place at birth at rest, and to determine if oscillations already relate to speech processing (during sentence presentation). Additionally, we could also explore if experience with the prenatally heard language already has an effect on neural oscillations at birth.

Furthermore, we were also interested in understanding how neural oscillations develop once infants had gained (postnatal) experience with spoken language. To investigate these developmental changes, we looked at 6-month-old infants (Chapter 3). We assessed how neural oscillations operate in the brain of 6-month-olds when they listen to sentences in their native language and in unfamiliar languages, as done with the newborns.

If neural oscillations subserve speech processing early in life, they could possibly be informative of later language development and its potential disorders or delays. Children are often not diagnosed with language disorder/delay until age 3 or beyond. However, disorders may be present in the brain or the auditory system from the beginning, with the behaviours allowing their diagnosis only detectable later in life. Neurophysiological measures may thus be useful tools for an earlier detection, providing the opportunity to perform early interventions. This thesis, therefore, explored whether neural oscillations could inform us about a child's later language capabilities. To do so, we followed our participants longitudinally after the neural oscillation measures at birth and at 6 months. Specifically, we collected information about their vocabulary size and their communicative skills at 12 and 18 months. We then assessed if the early neural measures could predict these later language outcomes (Chapter 4).

1.1 NEURAL OSCILLATIONS

Neurons can be thought of as oscillators; they fire action potentials repetitively at given frequencies (Wang, 2010). When large coalitions of neurons operate collectively in space and time (at scales of millimeters and milliseconds), they form neuronal assemblies (Badin et al., 2017). Neural oscillations correspond to the synchronous activity of neuronal assemblies. Neuronal assemblies can be coupled intrinsically, i.e. in the absence of stimulation (Giraud et al., 2007), forming a resting-state network, or by a common input, being part of the same functional network (Arnal et al., 2015).

Neural oscillations and their coupling have been claimed to play a role in many cognitive processes (Buzsaki 2006). They have been observed to be involved in memory, language, attention and sensorimotor integrations tasks (Wang, 2010). Most importantly for our

purposes, neural oscillations have been shown to play a role in speech and language processing (Bastiaansen & Hagoort, 2006; Bastiaansen et al., 2008; Giraud & Poeppel, 2012; Gross et al., 2013; Meyer, 2018).

Experimentally, neural oscillations can be evidenced by performing time-frequency analysis of electrophysiological recordings, some of which can be performed non-invasively on healthy participants (electroencephalography, EEG; magnetoencephalography, MEG), and some that due to their invasive nature are limited to patients undergoing surgical procedures for independent clinical reasons (stereotactic EEG, sEEG; electrocorticography, ECoG). In this dissertation we explored neural oscillations during rest and during speech processing by means of scalp EEG recordings.

1.1.1 Neural Oscillations in Animals

Neural oscillations have been extensively investigated in animals. Studies in awake and anesthetized animals have revealed that neural oscillations and their hierarchical organization are not specific to speech and language, and most importantly, not specific to humans. Experiments on cats and rats have shown that the amplitude of higher frequency oscillations appears to be modulated by the phase of lower frequency oscillations (Steriade et al., 1996; Lakatos et al., 2004; Lakatos et al., 2005). Furthermore, Lakatos et al. (2005) found that spontaneous neural oscillations in macaque monkeys are hierarchically organized: the phase of delta oscillations (1-4 Hz) modulates the amplitude of theta oscillations (4-10 Hz), and the phase of theta modulates the amplitude of gamma oscillations (30-50 Hz). These findings support the idea that neural oscillations have been evolutionarily available already in non-human animals, and that language may have evolved to match the brain's pre-existing oscillatory frequencies (Morillon et al., 2010; Meyer, 2018).

1.1.2 Neural Oscillations at Rest

Rest or resting state can be measured as an experimental condition where no stimulus is presented/delivered, and it is generally tested to capture the properties of intrinsic (spontaneous or endogenous) neural activity (Snyder & Raichle, 2012). In the past, spontaneous brain activity was regarded as noise, but experimental research demonstrated that it is not random. It is actually organized in sets of coherent functional networks (Deco et al., 2011). Electrophysiological recordings from the human auditory cortex have shown that oscillatory activity in the delta, theta, alpha and low-gamma bands may be observed during

resting state, as displayed in Fig. 1a (Giraud et al., 2007; Morillon et al., 2010; Giraud & Poeppel, 2012).

Resting state testing conditions are relatively unconstrained, and they can vary in numerous ways, from having eyes closed or open, with or without visual fixation (Snyder & Raichle, 2012), and were the presence of the subject's thinking is difficult to control for. Despite this, resting state networks exhibit high reproducibility (Damoiseaux et al., 2006) and moderate to high test-retest reliability (Franco et al., 2009; Allen et al., 2011).

However, resting state activity may be modified by prior stimulation, as it was shown for 3-month-old infants who exhibited higher functional connectivity at rest after 3 minutes of speech stimulation (Homae et al., 2011). This may be a reflection of the strengthening and/or reorganization of neural circuits as a result of stimulation, which are the neural markers of Hebbian learning. Here we investigated spontaneous neural oscillations at birth prior and after speech presentation with two aims: (1) providing a baseline for the language testing conditions, and (2) exploring if stimulation after-effects are already present at birth. Finding modifications in the resting state neural oscillations after speech stimulation is an indication that the neural mechanisms supporting learning and neural plasticity are already operational immediately at birth, thus paving the way for language development.

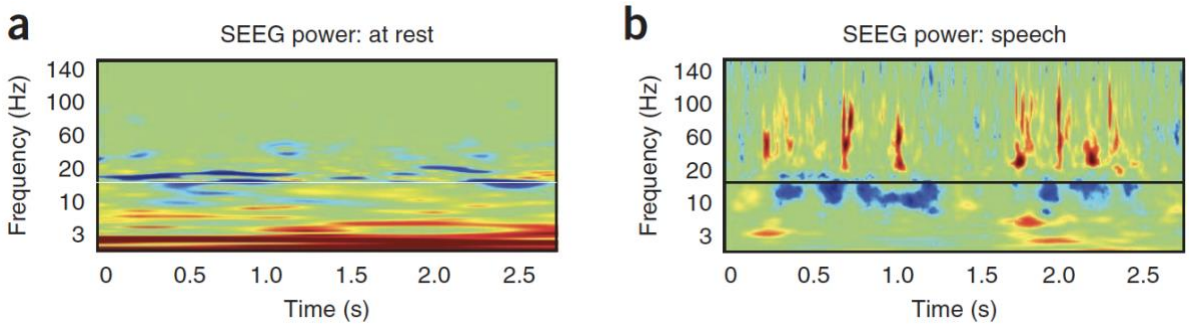


Figure 1. Time-frequency representations of adult cortical activity from intracortical recordings. **a)** During resting state. **b)** During speech processing. Figure taken from Giraud & Poeppel, 2012.

1.1.3 Neural Oscillations and Speech Processing

Speech processing is the set of cognitive, perceptual and neural processes that enable the segmentation and the identification of linguistically relevant phonological units in the speech signal (Meyer, 2018). During speech processing the auditory cortex is stimulated and resting state oscillations are transformed into temporally structured oscillations (Giraud & Poeppel, 2012), as illustrated in Fig. 1b. Giraud & Poeppel (2012) proposed that neural oscillations in the auditory cortex provide a first step in the process of deciphering continuous speech, by

providing the infrastructure to parse it and decode it. More specifically, they argue that neural oscillations in the low-gamma (25-35 Hz), theta (4-8 Hz) and delta (1-3 Hz) bands provide the brain with the means to convert the speech signal into phonemic (30-50 Hz), syllabic (4-7 Hz) and phrasal (1-2 Hz) units respectively, by aligning their phase to the salient points present in the speech signal at the onset or offset of phonemes, syllables and intonation phrases (Giraud & Poeppel, 2012; Meyer, 2018). This hypothesis is partly based on the close correspondence that exists between the frequencies at which these oscillations operate and the time scales of these speech units. Importantly, speech is organized into a hierarchy of phonological units, known as the prosodic hierarchy (Nespor & Vogel, 1986). Each smaller unit is embedded in a larger unit at the immediately higher level of the hierarchy. The largest unit is the utterance, which is made up of intonational phrases, which themselves are made up of phonological phrases, which in turn contain prosodic words, made up of syllables, constituted by phonemes. (The levels of the hierarchy may change as a function of the specificities of each language, e.g. in English syllables are organized into feet, while in Japanese, the level of morae is added between phonemes and syllables.) Neural oscillations are similarly organized into a hierarchy or a nesting/entrainment relation, whereby the lower frequency oscillations enslave the higher frequency ones, i.e. the phase/amplitude of the slower oscillations will determine or modulate the phase/amplitude of the faster oscillations. This hierarchical organization of the neural oscillations explains how we can process speech units of different size simultaneously, leading to the formulation of the multi-timescale model of speech perception (Giraud & Poeppel, 2012). The top-down coupling responsible for the nesting relation between oscillations has been empirically shown between theta and gamma oscillations (Morillon et al., 2012), and between delta and theta oscillations (Lakatos et al., 2005). It has been proposed that this top-down coupling could have two main purposes: (1) to ensure the temporal alignment of neuronal excitability with speech spectral information, and (2) to aid in the neural binding of discrete phonological units (intonation phrases, syllables, phonemes) (Meyer, 2018).

There is a growing body of experimental findings supporting the role of neural oscillations in speech processing and showing a link between neural oscillations and the processing of speech units.

Specifically:

- Gamma band oscillations have been suggested to aid in the discretization of syllables into (sub)phonemic units. Specifically, it has been shown that the amplitude

synchronization of low-gamma oscillations (Lehongre et al., 2011) to acoustic modulations at the phonemic rate (30 Hz) is related to phonemic-categorical perception. It is not yet clear, however, exactly what phonemic features are represented by gamma oscillations. This is partly due to methodological constrains, as gamma oscillations are not always well captured by non-invasive recording techniques (EEG, MEG).

- Theta band oscillations have been suggested to aid in the processing of syllables. This claim has mainly been based on two facts: (1) the syllabic rate of speech, roughly 4-5Hz (Ding et al. 2017, Varnet et al. 2017), corresponds to the frequencies of the theta band (Giraud & Poeppel, 2012), and (2) theta oscillations are phase synchronized across the experimental trials during speech processing (Luo & Poeppel, 2007; Howard & Poeppel, 2010). These latter results are based on across-trial phase coherence, and not the phase coherence between the auditory and the neural signals, making it hard to assess if the observed effect is really due to syllabic information (Meyer, 2018). Nevertheless, some more recent studies have performed a direct comparison between neural and acoustic signals (Gross et al., 2013; Vander Ghinst et al., 2016; Molinaro et al., 2016; Pefkou et al., 2017) and their results support the idea that there is indeed a relationship between theta oscillations and speech at the syllabic rate.
- Delta band oscillations have been claimed to process larger prosodic units such as intonational phrases. It has been shown that delta oscillations increase their phase coherence to the pitch contour of speech (Bourguignon et al., 2013; Mai et al., 2016; Molinaro et al., 2016), which in turn is related to intonational phrase boundaries (Frazier et al., 2006; Meyer, 2018).

1.1.4 Neural Oscillations and Language Comprehension

Not only does the speech stream need to be segmented into units, but the morphosyntactic organization and the meaning of these units also need to be decoded to achieve comprehension. Neural oscillations have been found to be involved in these processes underpinning syntactic and semantic processing and contributing to the prediction of upcoming linguistic material (Meyer, 2018).

Specifically, it has been suggested that neural oscillations in the delta, alpha and theta bands are associated to the following processes, respectively: (i) the grouping of words into syntactic phrases, (ii) the storage of these syntactic phrases in verbal working memory, and (iii) the retrieval of verbal information from working memory or from long-term memory for

the comprehension of the ongoing sentence (Meyer, 2018). Indeed, oscillations in the delta band have been shown to signal syntactic phrase boundaries (Ding et al., 2016); alpha oscillations have been associated with storage in working memory during sentence processing (Haarmann & Cameron, 2005); while theta band oscillations have been observed during memory retrieval (Bastiaansen & Hagoort, 2006).

Oscillations have been documented to underlie at least two aspects of predictions during sentence comprehension. Neural oscillations in the beta band have been associated with the semantic prediction of upcoming words in the ongoing sentence (Pefkou et al. 2017), whereas gamma oscillations may underlie the verification of prediction fit (Meyer, 2018, Lewis & Bastiaansen, 2015).

It is important to note that the two bodies of research that investigate how neural oscillations may subserve speech processing and language comprehension respectively have not been unified, and this remains an open issue. In Table 1 we present a summary of the neural oscillations that have been shown to be involved in each of these tasks.

Neural Oscillations	Speech Processing	Language Comprehension
Delta	Identification of intonation phrases	Formation of syntactic phrases
Theta	Identification of syllables	Retrieval from long-term memory
Alpha		Storage in working memory
Beta		Prediction
Gamma	Identification of subsyllabic units	Verification of prediction fit

Table 1. Summary of neural oscillations roles in speech processing and language comprehension.

1.2 SPEECH ENVELOPE TRACKING

A particular way in which neural oscillations attune to the auditory signal is known as envelope tracking. Speech envelope tracking is the phenomenon whereby neural activity in the auditory cortex synchronizes to the envelope of the incoming speech signal (Ahissar et al., 2001; Abrams et al., 2008; Kubanek et al., 2013). The mechanism of envelope tracking is similar to neural oscillations in that the auditory cortex synchronizes its activity with speech, but it involves a broader range of frequencies and a wider correspondence between speech and brain activity.

1.2.1 Definition of Speech Envelope

The speech envelope refers to the slow overall amplitude fluctuations that a speech signal has over time, which occur between about 2 and 50 Hz (Rosen, 1992). To illustrate this concept, we present in Fig. 2a the waveform of a speech signal (black curve) with its envelope (red curve) superimposed over it. Additionally in Fig. 2b, we display the speech envelope by itself.

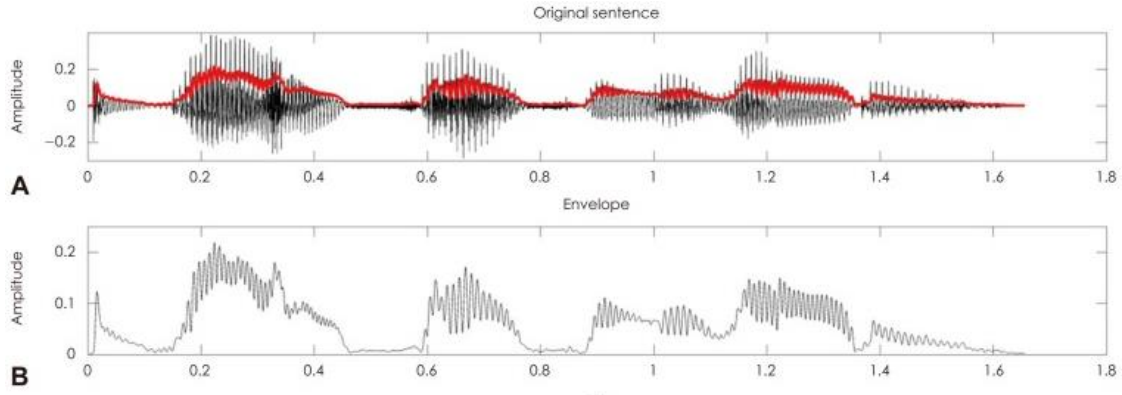


Figure 2. Time waveform of a speech signal. **a)** The envelope (red curve) is superimposed on the original speech signal. **b**) Plot of the broadband speech envelope. Figure taken from Moon & Hong, 2014.

The speech envelope conveys information about a variety of linguistic units. Figure 3 presents a list of the linguistic features that are carried by the speech envelope (Rosen, 1992).

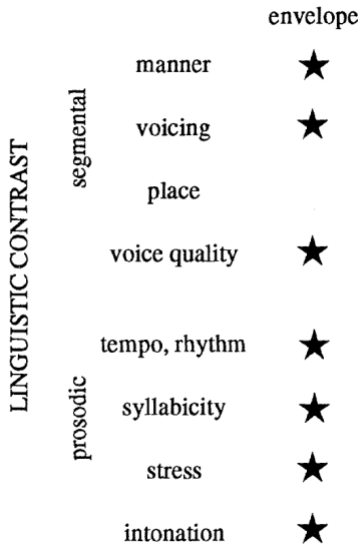


Figure 3. Linguistic features conveyed by the speech envelope (indicated by the presence of a star). Figure modified from Rosen, 1992.

1.2.2 Speech Envelope and Comprehension

The speech envelope has been argued to be crucial for speech perception. Behavioural studies have shown that comprehension is impaired when the speech envelope is suppressed (Drullman et al., 1994a; Drullman et al., 1994b), and that comprehension is nearly perfect when speech is degraded but the envelope is preserved, at least when presented in silence (Shannon et al., 1995). Electrophysiological studies have supported this claim by showing that speech envelope tracking deteriorates when comprehension decreases. Ahissar et al. (2001) found that when presented with accelerated speech the participant's level of comprehension was correlated with the phase-locking of their cortical activity to the speech envelope. Similarly, Peelle et al. (2013) showed that when presenting vocoded speech, the phase-locking between the neural activity and the speech envelope was enhanced when speech was intelligible.

Some other studies however, have questioned this proposal, as phase-locking to the speech envelope in the theta band, the oscillation whose frequency corresponds to the modulation frequency of the speech envelope (Ding et al., 2017; Varnet et al., 2017), has been found to be independent of comprehension. Brain responses in the theta band have been shown to track the speech envelope even when speech is time-compressed at a rate that renders it incomprehensible to adult listeners (Pefkou et al., 2017). Moreover, normal sentences, readily comprehensible to listeners, have been shown to evoke similar phase-encoding responses as their unintelligible, time-reversed counterparts (Howard & Poeppel, 2010); and stronger phase-locking results have been found in adults when listening to sentences in their second language than in their native language (Song & Iverson, 2018). These results thus call into question whether the occurrence of envelope tracking is sufficient for comprehension to take place. More generally, it remains controversial whether speech comprehension and envelope tracking are causally linked.

To shed new light on this debate, here we investigated speech envelope tracking at birth and at 6 months of age. By assessing speech envelope tracking in familiar and unfamiliar languages, we can address: (i) whether envelope tracking is a neural mechanism involved in language comprehension, (ii) whether general experience with speech is necessary for it to emerge, and (iii) whether familiarity with the native language modulates it.

1.2.3 Measures of Envelope Tracking

When assessing speech envelope tracking, a variety of measures have been proposed in the literature. Some of them include: cross-correlation (Abrams et al., 2008), inter-trial phase coherence (Luo & Poeppel, 2007), inter-trial power coherence (Luo & Poeppel, 2007), phase-coherence (Peelle et al., 2013; Pefkou et al., 2017), Cross-Spectral Density (Pefkou et al., 2017), and multivariate temporal response functions (mTRFs) (Kalashnikova et al., 2018). In our study we focused on two measures because they are the most explicitly developed ones in the literature: cross-correlation, assessing how the brain encodes the amplitude of the speech envelope, and phase-coherence, assessing how the brain encodes its phase.

Amplitude synchronization occurs when the amplitude of the neural activity follows the contours of the speech envelope, and is often assessed using cross-correlation analysis (Abrams et al., 2008; Nourski et al., 2009; Kubanek et al., 2013). Phase synchronization occurs when the auditory cortex modulates the phase of its ongoing oscillations to match the phase of the speech envelope (i.e. phase-locking), and is often assessed by computing the phase-coherence between the neural oscillations and the acoustic signal (Peelle et al., 2013; Pefkou et al., 2017). For illustration, Fig. 4 shows how EEG recordings from school children (red and blue curves) track the speech envelope (black curve) of a sentence in their native language both in amplitude and phase.

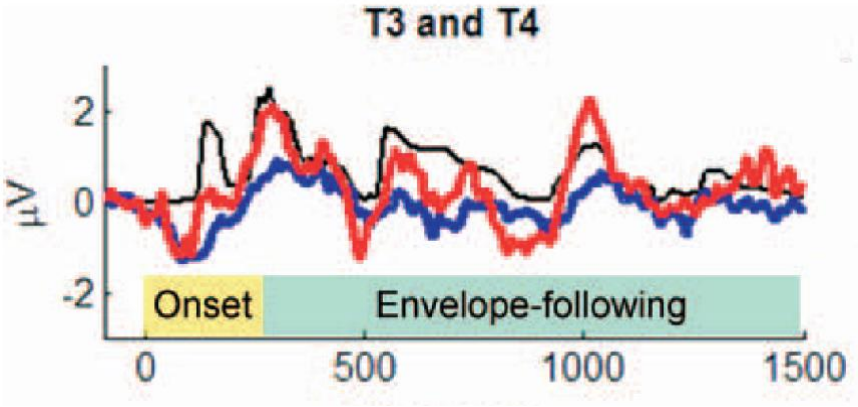


Figure 4. Speech envelope tracking in children (9-13 years old). Grand average cortical responses from EEG recordings during speech processing. The black curve represents the broadband speech envelope; the red curve represents cortical activity measured at electrode T4 (right-hemisphere, also called T8); the blue curve represents cortical activity measured at electrode T3 (left-hemisphere, also called T7). Figure taken from Abrams et al., 2008.

Studies in the literature also differ in the frequency ranges of the neural data they take into account to assess envelope tracking. Studies evaluating amplitude synchronization have

found envelope tracking in a wide range of frequencies, differing between non-invasive (*1-40 Hz* in Abrams et al., 2008) and invasive recordings (*70-250 Hz* in Nourski et al., 2009; *75-115 Hz* in Kubanek et al., 2013). For the studies evaluating phase synchronization there is more convergence in the frequency ranges in which envelope tracking is found between non-invasive (*4-8 Hz* in Luo & Poeppel, 2007; *3-7 Hz* in Howard & Poeppel, 2010; *4-7 Hz* in Peelle et al., 2013) and invasive recordings (*1-7 Hz* in Golumbic et al., 2013), mainly situated in the slow frequencies (delta and mainly theta range). The difference in the frequency ranges in which amplitude and phase synchronization have been found can be explained by the nesting relationship in neural oscillations, in which the phase of slower oscillations modulates the amplitude of the faster ones. This explains why when synchronizing to an external signal which is slow (i.e. the speech envelope), the slow oscillations synchronize in phase, as their frequencies are close to the frequency of the envelope, and the fast oscillations synchronize in amplitude, as their frequencies are too different from the frequency of the envelope to allow for synching in frequency/phase. Thus the amplitude of the higher frequencies (and their synchronization) is modulated by the phase of the slower frequencies. Taking into account that invasive recordings are more sensitive to high frequencies, it not surprising that studies investigating amplitude synchronization invasively include higher frequency ranges than non-invasive studies.

It is important to note that the relationship between theta oscillations and syllabic information that we described earlier in this chapter (in the *Neural Oscillations and Speech Processing* section) is similar to the same speech envelope tracking by means of phase synchronization described here. In other words, the relationship between theta oscillations and syllabic information has been proposed to be based on the phase synchrony found between the neural activity (in the theta band) and the speech envelope.

In the literature only very few studies have assessed amplitude and phase synchronization at the same time (Golumbic et al., 2013). Therefore, it is not clear whether speech processing and language comprehension require both low-frequency phase and high-frequency amplitude synchronization to take place. In our study we investigated speech envelope tracking using both measures to explore if they co-occur in infancy.

1.3 SPEECH PERCEPTION AND LANGUAGE PROCESSING DURING INFANCY

Despite unresolved questions, in the light of the evidence discussed above, it is now uncontroversial that neural oscillatory activity plays a role in speech perception and language processing in adults. But do these same neural mechanisms support language acquisition?

To answer this question, as part of this dissertation, we have followed up a cohort of the infants who participated in the two EEG sessions longitudinally in order to assess their language outcome during the second year of life. We have thus collected data from our participants at four important stages of development: birth (EEG), 6 months (EEG), 12 months (vocabulary size, MacArthur-Bates CDI) and 18 months (vocabulary size, MacArthur-Bates CDI).

In order to better understand what aspects of language development neural oscillations may subserve, we now review the major milestones of speech and language processing in infancy, with special attention to aspects that may be related to neural oscillations at the ages investigated.

1.3.1 Prenatal Hearing

The human cochlea is operational from around the 20th week of gestation, when it starts sending auditory messages to the brain, and reaches almost adult-like maturity at around 30 weeks. However, this is not the case of the auditory brain structures, which do not reach full maturity during gestation, and continue developing after birth, well into childhood (Pujol & Lavigne-Rebillard, 1991).

There is evidence that human foetuses are able to hear in utero (Granier-Deferre et al., 1985), and that maternal speech and heartbeats are audible in the womb (Querlue & Renard, 1981; Querlue et al., 1981; Walker et al., 1971; DeCasper & Spence, 1986). However, the speech that reaches foetuses in the womb is not like the speech we experience in our daily life. As speech travels to the womb it gets attenuated by the maternal tissues, which act as a low-pass filter, exhibiting higher attenuation for higher frequencies (Pujol & Lavigne-Rebillard, 1991). This filtering effect suppresses the fine details that characterize individual speech sounds. However, there is good transmission of low-frequency information, and thanks to this some low-frequency properties of speech like rhythm, melody, and intonation are preserved and reach the womb (Vince & Armitage, 1980; Gerhardt et al., 1990; Lecanuet & Schaal, 2002; Gervain & Werker, 2008). The main speech input that foetuses receive is that

of their mother's voice. Non-maternal speech is highly attenuated by the maternal tissues and/or masked by the intrauterine sounds (Querlue & Renard, 1981; Querlue et al., 1981; DeCasper & Spence, 1986). This exposure to the maternal voice gives them access to the maternal language and its linguistic information.

1.3.2 Speech Perception at Birth

Thanks to a great body of research comprising behavioural and imaging studies, we know how sophisticated newborn babies' perceptual and learning abilities are. At birth they are equipped with a rich set of speech perception abilities that help them acquire language. Some of these abilities are universal, in place independently of their prenatal experience; while others are shaped by the speech they heard in utero (Ortiz-Barajas & Gervain, submitted).

Universal Abilities

We consider an ability universal, i.e. the experience-independent at birth, if prenatal experience with speech heard in utero doesn't modulate it. Thus newborns recognize human speech and prefer it over equally complex speech analogues (Vouloumanos & Werker, 2007). They are also able to discriminate two languages that belong to different rhythmic classes (Mehler et al. 1988; Ramus et al. 2000), even if they are unfamiliar (Nazzi et al. 1998). It has also been shown that newborns can detect the acoustic cues that signal word boundaries (Christophe et al. 1994), and that they can discriminate words with different patterns of lexical stress (stress-initial vs. stress-final) (Sansavini et al. 1997). At birth, infants possess a (close to) universal phoneme discrimination ability, as they are able to discriminate almost all the phonemes present in all the languages (Eimas et al., 1971; Dehaene-Lambertz & Baillet, 1998). Additionally, it has been shown that newborns can easily detect and learn repetition-based patterns (e.g. "mu ba ba", "pe na na") and discriminate them from similar random sequences (e.g. "mu ba ge") (Gervain et al., 2008; Gervain et al., 2012). Besides possessing this wide range of universal abilities, newborns have also been found to exhibit an adult-like brain organization for speech, with a network, which is at least partly left lateralized, and involves superior temporal and inferior frontal regions (Peña et al., 2003; Mahmoudzadeh et al., 2013).

Abilities Shaped by Prenatal Experience

Some speech perception abilities that newborns possess have been shown to depend on the speech heard in utero. From having heard their mother's voice in utero, newborns recognize it and prefer it over other female voices (DeCasper & Fifer, 1980). They also recognize and prefer the language heard in utero (the native language) over other languages (Mehler et al. 1988; Moon et al. 1993). Furthermore, newborns are able to recognize some of the vowels (Moon et al., 2013), and the prosodic patterns of their native language (Abboub et al., 2016), and can remember a short speech passage heard frequently in utero during the last weeks of gestation (DeCasper & Spence, 1986). In line with this prenatal learning, newborns have been found to exhibit brain specialization (left lateralization and/or increase activation) for the prenatally heard language (Peña et al., 2003; May et al. 2011; Sato et al., 2012).

1.3.3 Speech Perception During the First Year of Life

Even though the human cochlea is fully mature at birth, the cortical auditory structures are not, and they continue to develop during the first years of life (until 4-8 years of age), forming connections depending on the stimulation they receive (Pujol & Lavigne-Rebillard, 1991).

As it was previously mentioned, the speech signal that reaches the womb is low-pass filtered by the maternal tissues. Therefore, it is only after birth that infants encounter full-band speech for the first time. As a consequence of the interplay between post-natal speech experience and the continuous development of the auditory cortex, new speech perception abilities emerge during the first years of life. By about 4 months of age, infants start to discriminate their native language from a rhythmically similar one. Sometime between 4 and 6 months infants start to segment out word forms and to associate them with some basic meaning; and by the end of the first year of life the infant brain is attuned to the native language. Below, we review these abilities in detail.

Rhythmic Discrimination of Languages

Speech rhythm is defined as the temporal organization of prominence that contributes to the prosodic structure of utterances (Astésano, 2001 as cited by Magne et al., 2004). Linguists have classified the languages of the world in three categories claiming that they represent their different rhythms: syllable-timed, stress-timed, and mora-timed (Ramus et al., 1999; Gervain & Werker, 2008). Each category has an element that reoccurs at regular time

intervals and establishes a temporal organization. ***Syllable-timed*** languages are those for which the syllable was claimed to be the organizing time unit, and include French, Spanish and Italian. In syllable-timed languages each syllable has equal stress and this gives them what has been called a machine-gun rhythm. ***Stress-timed*** languages are those for which the interstress interval was claimed to be the dominant feature of the rhythmic timing, and include English, German, Dutch and Arabic. In stress-timed languages some syllables are longer than others, but those that are stressed have approximately the same length, and this gives them what has been described as a Morse code rhythm (Denham & Lobeck, 2012). ***Mora-timed*** languages are those for which the reoccurring unit is believed to be the mora, and include Japanese and Tamil (Ramus et al., 1999; Gervain & Werker, 2008; Ladefoged & Johnson, 2011).

These rhythmic differences are perceived easily by any of us, however this is not surprising considering that we have a vast knowledge of language. What is most interesting is that newborn babies can also perceive these rhythmic differences shortly after birth (Mehler et al. 1988; Ramus et al., 2000), and that they are able to do it even if the languages being contrasted are unfamiliar to them (non prenatally heard) (Nazzi et al., 1998). More importantly, monkeys have also been shown to detect these rhythmic differences, therefore this rhythmic discrimination ability is not a linguistic, but rather a general auditory ability (Ramus et al., 2000). How this language discrimination takes place at birth at the neural level is not fully understood. It has been argued that newborns might rely on suprasegmental, prosodic cues, as they are able to do it even when the speech stimuli have been low-pass filtered (which removes the segmental information) (Mehler et al., 1988; Gervain & Werker, 2008). One of the prosodic cues that seems to play an important role is rhythm. Ramus et al. (1999) has operationally defined rhythm as the relative proportion (and variability) of vowels and consonants in the speech signal. Since then, other acoustic metrics of rhythm have also been proposed (e.g. Dellwo, 2006; Grabe & Low, 2002). Furthermore, we have shown (Varnet et al., 2017) that languages belonging to different rhythmic classes show different modulation rates in their amplitude modulation spectra, with syllable-timed languages having somewhat faster modulations than stress-timed languages. Newborns seem to depend on some of these acoustic correlates of linguistic rhythm to perform language discrimination. As a consequence, they are unable to tell apart two languages that belong to the same rhythmic category (Nazzi et al., 1998).

Being able to differentiate two languages from the same rhythmic category is not an easy task. Infants achieve this within-rhythmic class discrimination ability approximately 4 months after birth (for full-term babies, and later for premature babies, Peña et al., 2010), and only if they have had postnatal exposure to at least one of the languages in question (Bosch & Sebastián-Gallés, 1997; Molnar et al., 2013). This suggests that this challenging task requires a more mature brain, and knowledge of the specific sound patterns (e.g. segmental cues) of the languages (Peña et al., 2010; Molnar et al., 2013).

Most of the studies investigating rhythmic discrimination were conducted behaviourally, but Peña et al. (2010) have investigated the oscillatory signature of discrimination in the gamma band. They have shown that during language discrimination, 6-month-old infants exhibit higher gamma activation for their native language (Spanish) than for an unfamiliar language from the same rhythmic class (Italian), as well as for an unfamiliar language from a different rhythmic class (Japanese). This latter language already triggered lower gamma power at 3 months, the youngest age measured, than Spanish and French. This finding has provided insight into the neural mechanisms underling language discrimination. However, little is known about the neural mechanisms that support rhythmic discrimination at birth. In our study we are investigating the behaviour of neural oscillations at birth for familiar and unfamiliar languages from different rhythmic classes. This will allow us to gain insight about the relationship between neural oscillations and language discrimination.

Segmentation and Word Learning

Identifying words from a speech stream is indispensable for building a vocabulary. For infants to identify single words in continuous speech, they need to be able to segment it. It has been shown that around 7 months of age infants start to detect familiar words in speech (Jusczyk & Aslin, 1995). There are various cues that infants can rely on to identify words. The first one is the stress patterns typical of their native language, which they have been shown to be sensitive to between 6 and 9 months of age (Jusczyk et al., 1993a). The second one is phonotactic, which allows them to identify what phoneme combinations are frequent at the beginning/end of a word, and which has been found to help 9-month-olds during word segmentation (Saffran & Thiessen, 2003). The third one is the position of allophones within words, which has been shown to be sufficient to signal word boundaries in 10-month-olds (Jusczyk et al., 1999b). The fourth cue is statistical regularities which inform them of the

probability of having syllables co-occur in the same word, and which has been shown to help already at 8 months of age (Saffran et al., 1996).

Segmentation abilities necessary for the beginning of word learning are thus in place between 6-9 months of age. It is thus not surprising that by this same age, infants have been shown to start associating possible word forms to possible referents/meanings, at least for a handful of very frequently occurring words such as body parts or food items (Tincoff & Jusczyk, 1999, 2012; Bergelson & Swingley, 2012).

Perceptual Attunement to the Native Language

Languages have different phoneme repertoires. In English for example the phonetic units 'r' and 'l' are differentiated, while in Japanese they are combined into a single phonemic category (Kuhl, 2004; Gervain & Werker, 2008). At birth infants have a (close to) universal perceptual ability and they are able to distinguish between almost all the phonemes present in the languages of the world (Eimas et al., 1971; Dehaene-Lambertz & Baillet, 1998). During the second half of the first year of life, this universal ability transitions into a more language-specific one (Werker & Tees, 1984; Kuhl, 2004). While at 6 months of age, infants can easily discriminate phonetic units that are not used in the language they are learning, by the time they reach 12 months this discrimination becomes difficult to perform (Werker & Tees, 1984; Best & McRoberts, 2003). Parallel to this non-native perceptual decline, the ability to discriminate the phonetic units of their native language seems to improve by the end of the first year (Cheour et al., 1998; Rivera-Gaxiola et al., 2005b).

Several hypotheses have been proposed to explain this perceptual attunement to the native language. While the accounts differ in their details, most of them argue that infants rely on the statistical distributions of the sounds of their native language to highlight the phonetic units that are used in their language and to de-emphasize those that are not (Maye et al., 2002; Kuhl, 2004; Gervain & Werker, 2008). At the neural level, this has been linked to the critical period for language in infancy when neural sensitivity to the speech contrasts is preserved upon exposure, and lost without it (Weikum et al., 2012; Best et al., 2016).

In sum, during the first year of life, infants pass a number of important milestones in speech perception and language development, the most relevant ones of which have been described above. Importantly, some of these are taking place in parallel and possibly mutually

supporting each other. Figure 5 presents a timeline that illustrates how the experience with the speech signal and the perceptual abilities change from gestation to the end of the first year of life.

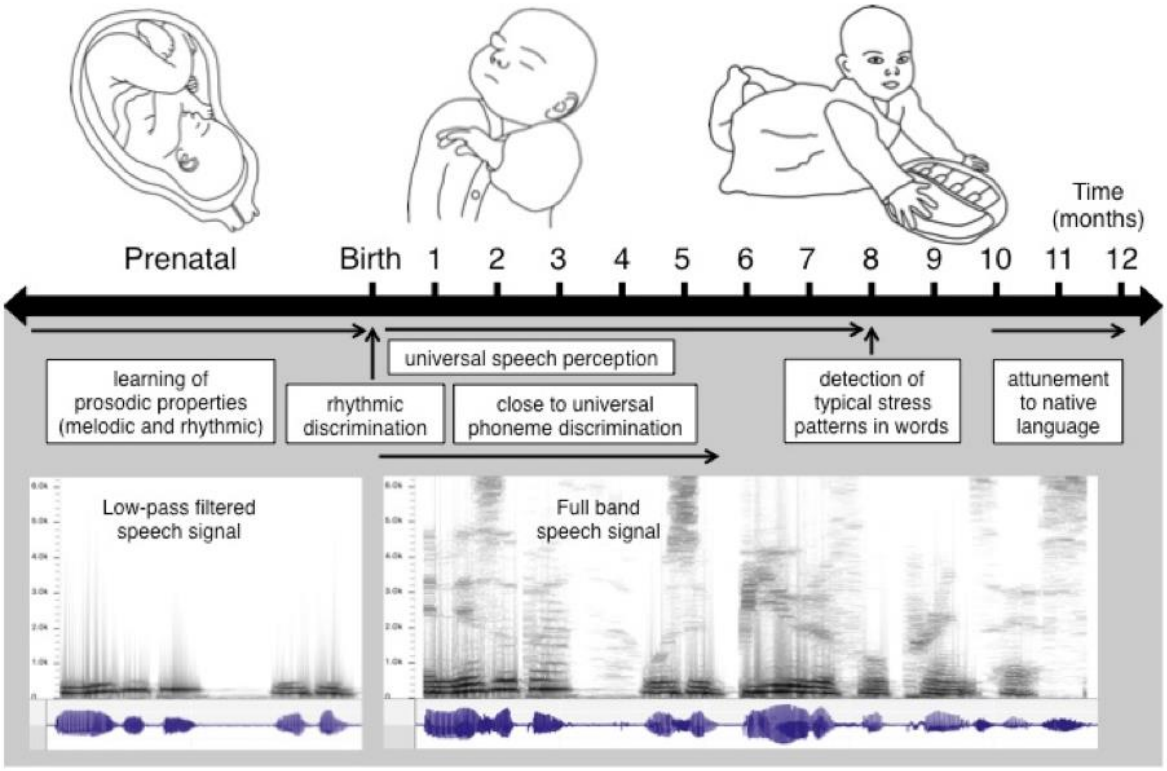


Figure 5. Timeline of infants' changing experience with the speech signal, and their developing speech perception abilities. Figure taken from Ortiz-Barajas & Gervain (submitted).

Our study will shed light on whether some of these abilities may be explained, at the neural level, by the emergence or refinement of neural oscillations, providing evidence for a role of this mechanism in language development.

1.4 PREDICTING LANGUAGE OUTCOME

The role of neural oscillations in language development will be further investigated by exploring whether they could predict later language outcome. Most developmental studies investigating speech perception abilities have based their conclusions on the general trends observed at the group level. However, in the last two decades a growing body of research has started to pay attention to the individual variability in the behavioural and neurophysiological measures of speech perception abilities during the first year of life, and have tried to link it to later measures of language development during the second and third year of life (Cristia et al., 2014). Looking at these individual differences as possible predictors of language outcome is

relevant for identifying early cues of language deficiencies. The importance of finding speech perception measures that could predict later language development lies on the fact that children with learning or reading disabilities typically show deficits in speech perception already earlier than when their disorder is diagnosed (Kuhl et al., 2005). Currently, language impairments are difficult to diagnose before the age of 3-4 years (Cristia et al., 2014), therefore identifying measures that could allow their earlier detection is fundamental for the design of earlier interventions. If neural oscillations are fundamental for speech perception already at the beginning of human development, then they may serve as predictors of later language outcomes.

Tsao and colleagues (Tsao et al., 2004) performed one of the first studies exploring the link between speech perception and language acquisition before the age of 2 years. They used the conditioned head-turn task to test 6-month-old infants on a speech discrimination task (a vowel contrast perceived by adults as native), and assessed their language development during the second year of life using the MacArthur Communicative Development Inventory (CDI) at 13, 16, and 24 months. They found significant correlations between the speech perception skills at 6 months and vocabulary measures of comprehension at 13 months and production at 24 months. In a follow-up study, Kuhl and colleagues (Kuhl et al., 2005) tested a similar paradigm on 7 month-olds, but this time with two conditions: one contrast from their native language, and one from a non-native language. They found that both native and non-native phonetic perception abilities were related to later measures of language outcome but in opposite directions: better native-language discrimination at 7 months was positively correlated to expressive vocabulary at 18 and 24 months, whereas better non-native-language discrimination was negatively correlated to expressive vocabulary at 18 and 24 months. These findings were supported by an electrophysiological study comparing ERP responses in 11-month-olds to native and non-native speech contrasts (Rivera-Gaxiola et al., 2005a). They showed that infants who exhibited a larger (more positive) amplitude for a given ERP component produced more words at 18, 22, 25, 27 and 30 months, than those who had a more negative peak. A later ERP study from the same team showed that ERP responses to native and non-native contrasts at 7 months also related to later language outcomes, again in opposing directions (Kuhl et al., 2008). They suggested that increased sensitivity in the perception of native phonetic contrasts is indicative of neural commitment to the native language, and the sensitivity to non-native contrasts reveals uncommitted neural circuitry (Kuhl et al., 2008). The ERP responses then seem to be a reflection of this level of neural

commitment, which in turn predicts language scores at later ages (Rivera-Gaxiola et al., 2005a).

Other studies assessing different speech perception abilities have also found predictors of later language measures. Cristia & Seidl (2011) showed that prosodic sensitivity at 6 months predicts vocabulary size at 24 months, as measured with the MacArthur-Bates Communicative Developmental Inventory (CDI). Prosodic sensitivity was tested with the Head-turn Preference Procedure (HPP) where infants were presented with two versions of the same sequence of words, but where only one was uttered as a well-formed prosodic unit. Those infants who preferred the well-formed intonational units had a larger receptive and productive vocabulary score at 24 months, compared to those who did not express this preference pattern. More Recently, Hoareau (2019) investigated whether word segmentation abilities by means of statistical mechanisms at the age of 8 months could predict later language outcome. She found that infants who exhibited better segmentation skills at 8 months had higher expressive vocabulary scores by the time they were 12 months.

1.5 RESEARCH QUESTIONS AND HYPOTHESES

In the light of the above, the current thesis seeks to investigate whether neural oscillations are already in place at birth and whether they relate to speech processing. Based on the fact that neural oscillations are a neural mechanism not exclusive to humans or language, we hypothesize that they are present in the human brain at birth. Additionally, considering that newborns exhibit an adult-like brain specialization for speech and express a wide range of speech perception abilities, we do expect to find neural oscillations to be related to speech processing at birth. Specifically, if prenatal prosody plays an important role, then we expect low frequency oscillations (delta, theta) to be present. We also expect no beta oscillations as they vehicle top-down grammatical and vocabulary knowledge (Pefkou et al, 2017).

We also seek to investigate whether the synchronization of neural oscillations to the speech envelope (i.e. speech envelope tracking) takes place at birth or whether it requires extended experience with speech to emerge. Based on the studies arguing a relationship between speech envelope tracking and comprehension, we hypothesize that this phenomenon might not be present at birth, or at least not for the unfamiliar languages (not heard prenatally). However, based on the studies showing phase synchronization to the envelope of unintelligible speech, we hypothesize that envelope tracking by means of phase synchronization could be present at birth.

We are also interested in exploring how neural oscillations and their relationship with speech processing change during the first 6 months of life. Based on previous findings (Peña et al., 2010), we expect to find developmental differences in gamma oscillations between birth and 6 months. More specifically, we expect gamma activity at 6 months to be stronger for the native language (French) than for the rhythmically similar one (Spanish), reflecting the within-rhythmic class language discrimination that they have been shown to express at this age.

Furthermore, we are also interested in investigating how speech envelope tracking changes during the first 6 months of life. We expect that postnatal exposure to speech in general (non specific to any language) would have an effect on speech envelope tracking, and that we would find tracking (even weakly) in the three languages. More specifically, we expect that experience with the native language would be translated into enhanced tracking of the native language.

Lastly, we aim to explore whether neural oscillations during the first months of life relate to language development during the second year. We expect that neural oscillations at 6 months could represent better candidates for predictors than neural oscillations at birth based on them potentially reflecting speech perception and learning abilities. However, we have no a priori hypothesis about which neural oscillations to use as predictors as we first ought to determine which oscillations reflect speech processing at each age (birth and 6 months).

To address these questions, we tested infants' neural oscillatory response to the native language (French), an unfamiliar, but rhythmically similar language (Spanish) and an unfamiliar and rhythmically different language (English), using EEG. We tested infants at birth, to be able to explore the role of prenatal experience and the emergence of oscillations, at 6 months, reflecting the impact of postnatal experience, and we followed up their vocabulary size at 12 and 18 months using the MacArthur-Bates Communicative Developmental Inventory in order to track their language development and relate it to their neural activity measured at birth and 6 months.

CHAPTER 2: Neural oscillations in the human newborn brain

2.1 INTRODUCTION

Neural oscillations in adults have been proposed to be involved in speech processing, by providing the brain with the means to convert the speech signal into different linguistic units at several levels at the same time (Giraud & Poeppel, 2012). However, little is known about how neural oscillations may subserve speech processing in the developing human brain. Neural oscillations in the infant brain have been shown to be involved in language discrimination (Peña et al., 2010): at 3 months, infants exhibit higher gamma activation for their native language (Spanish) than for a rhythmically different unfamiliar language (Japanese), reflecting the rhythmic discrimination ability that is in place from birth. Furthermore, by 6 months infants exhibit higher gamma activation for their native language (Spanish) than for a rhythmically similar unfamiliar language (Italian), a difference that was not observed at 3 months, and that reflects the within-rhythmic class discrimination ability that infants attain at about 4 months of life.

Up to date, it remains unknown whether neural oscillations' attunement to speech comes from extended experience with spoken language, or if it is a basic feature of the human auditory system. To shed some light on this developing field, we investigated neural oscillations and their relationship with speech processing at the very beginning of life with the aim to answer the following questions:

i. Are neural oscillations in place at birth?
ii. Do neural oscillations relate to speech processing at birth?
iii. Does familiarity with the prenatally heard language have an effect on the behaviour of neural oscillations?
iv. Do neural oscillations relate to the ability newborns have to discriminate rhythmically different languages?
v. Is there synchronization between neural oscillations and the speech envelope (i.e. speech envelope tracking) at birth?
vi. Is envelope tracking at birth shaped by prenatal experience?

To tackle these developmental questions, we tested newborns in five conditions (Fig. 6) while recording their neural activity with EEG: (1) at rest, before any speech stimulus was presented, (2) while passively listening to utterances in their native language (French), (3) a rhythmically similar unfamiliar language (Spanish), (4) and in a rhythmically different unfamiliar language (English), as well as (5) again at rest, after all the auditory stimulation

had taken place. All the conditions were part of one single experiment, and each participant heard all five conditions.

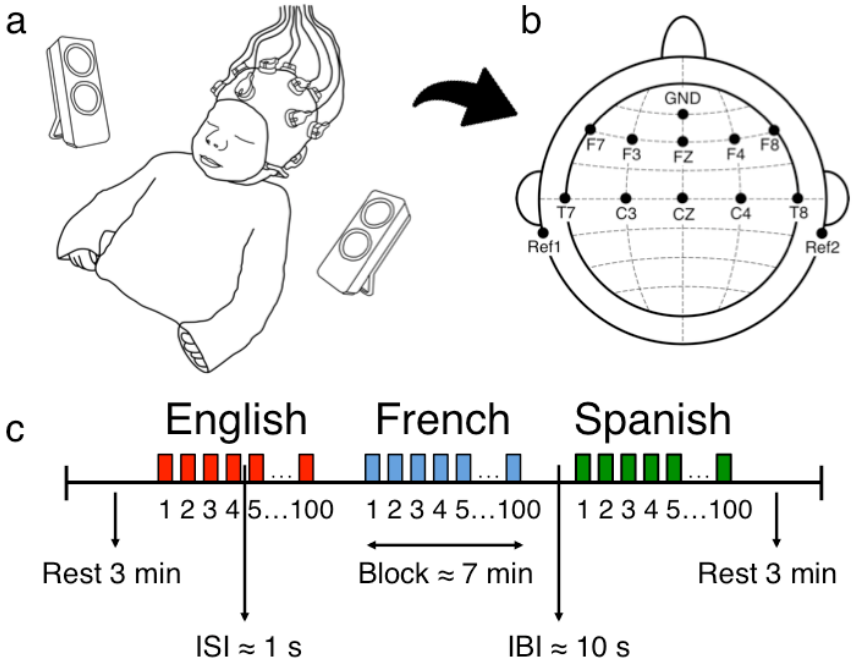


Figure 6. EEG experimental setup and design. **a)** Newborn with EEG cap. **b)** Location of recorded channels according to the international 10-20 system. **c)** Experiment block design. ISI: Interstimulus interval, IBI: Interblock interval.

To ensure that participants had very little post-natal experience with speech, we tested newborns within their first 5 days of life. This means that their experience with speech was mostly prenatal. To explore the effect of language familiarity as a result of prenatal exposure, we tested newborns who were born to French monolingual mothers, therefore, their prenatal experience came only from this language. We presented them with sentences in their native language (French), a rhythmically similar unfamiliar language (Spanish), and in a rhythmically different unfamiliar language (English). At birth, infants are able to discriminate rhythmically different languages, but unable to discriminate rhythmically similar ones (Nazzi et al., 1998; Ramus et al., 2000). Therefore, in our experiment newborns could discriminate French from English, and Spanish from English, but not French from Spanish. By testing these three language conditions, we were able to explore the relationship between neural oscillations and language discrimination at birth. Additionally, we tested resting state prior and after speech presentation to explore if stimulation after-effects are already present at birth.

Even though the relationship between neural oscillations and speech processing has not been broadly studied during development, we have some hypotheses of what we might find on this exploratory research: (1) if prenatal prosody plays an important role during speech processing in the newborn brain, then we expect low frequency neural oscillations (delta and theta) to be present at birth. (2) Additionally, if gamma oscillations are present, we expect them to have higher amplitudes when processing the native language (French) and the rhythmically similar unfamiliar language (Spanish) than for the rhythmically different unfamiliar one (English) (Peña et. al., 2010). (3) We also expect no activity in the beta band as it vehicles top-down grammatical and vocabulary knowledge (Pefkou et al., 2017).

Presenting newborns with the familiar (French) and unfamiliar (Spanish and English) language conditions also allowed us to explore if speech envelope tracking takes place at birth. We hypothesize that if envelope tracking represents a basic auditory mechanism, then we might find it in the three languages. Alternatively, if envelope tracking is related to comprehension then we should not find it at birth. Additionally, we were able to explore whether envelope tracking is shaped by prenatal experience. If we find envelope tracking to be similar for French and Spanish, and different to English, it means that it is shaped by the language heard in utero. If instead we find that tracking is similar for the three languages, then it means it is not modulated by language familiarity.

In order to address the various questions of our study, we analysed the collected EEG data applying different methodologies for each group of questions: time-frequency analysis for neural oscillations; and cross-correlation and phase-coherence for speech envelope tracking (more details in the methods section).

2.2 MATERIALS AND METHODS

2.2.1 Participants

Newborn participants were recruited at the maternity ward of the Robert-Debré Hospital in Paris, and were tested during their hospital stay. All parents gave written informed consent prior to participation, and were present during the testing session. The protocol for this study was approved by the CER Paris Descartes ethics committee of the Paris Descartes University. The inclusion criteria were: i) being full-term and healthy, ii) having a birth weight > 2800 g, iii) having an Apgar score > 8, iv) being maximum 5 days old, and v) being born to French native speaker mothers who spoke this language at least 80% of the time during the last trimester of the pregnancy according to self-report. A total of 55 full-term, healthy newborns

(age 2.38 ± 1.13 d; range 1-5 d; 22 girls, 33 boys) were tested. Eight participants were excluded from data analysis due to: bad data quality resulting in an insufficient number of non-rejected trials (n = 4), not finishing the experiment due to fussiness and crying (n = 3), or technical problems (n = 1). Thus, electrophysiological data from 47 newborns (age 2.45 ± 1.18 d; range 1-5 d; 20 girls, 27 boys) were included in the analysis.

2.2.2 Procedure

We tested newborns by presenting them with naturally spoken sentences in three languages while simultaneously recording their neural activity using electroencephalography (EEG) (Fig. 6a). The EEG recordings were conducted in a dimmed, quiet room at the Robert-Debré Hospital in Paris. Participants were divided into 3 groups, and each group heard one of three possible sets of sentences: 14 newborns heard set1, 16 newborns heard set2, and 17 newborns heard set3. During the recording session, newborns were comfortably asleep or at rest in their hospital bassinets. The stimuli were delivered bilaterally through two loudspeakers positioned on each side of the bassinet using the experimental software E-Prime. The sound volume was set to a comfortable conversational level (~65-70 dB).

Following the procedure of a previous study investigating envelope tracking in older children (Abrams et al., 2008), we presented participants with one sentence per language, and repeated it 100 times to ensure sufficiently good data quality. The experiment consisted of 3 blocks, each block containing the 100 repetitions of the test sentence in a given language, each block thus lasted around 7 minutes. An interstimulus interval of random duration (between 1 – 1.5 s) was introduced between sentence repetitions, and an interblock interval of 10 s was introduced between language blocks (Fig. 6c). The order of the languages was randomized and approximately counter-balanced across participants. Additionally, we recorded 3 minutes of resting state at the beginning of the experiment, before any stimulation took place, plus 3 minutes at the end, after all the sentence repetitions were over. The entire recording session lasted about 27 minutes.

2.2.3 Stimuli

We tested newborns in the following three languages: their native language (French), a rhythmically similar unfamiliar language (Spanish), and a rhythmically different unfamiliar language (English). The stimuli consisted of sentences taken from the story Goldilocks and the Three Bears. Three sets of sentences were used, where each set comprised the translation of a single utterance into the 3 languages (English, French and Spanish). The translations

were slightly modified in order to match sentence duration across languages within the same set. All sentences were recorded in mild infant-directed speech by a female native speaker of each language (a different speaker for each language), at a sampling rate of 44.1 kHz. The intensity of all recordings was adjusted to 77 dB. The three sets of sentences used in this study, as well as their duration and syllabic rate details are listed in Table 2. The speech envelope of each corresponding sentence is displayed in Fig. 7.

	Set 1	**Set 2**	**Set 3**
English	The bears lived all together in a beautiful house 2,50 s / 5 Hz	The bears decided to take a walk that day 2,16 s / 5 Hz	The little bears saw that their chairs were used by the naughty girl 2,96 s / 4 Hz
French	Les ours habitaient tous ensemble dans une maison 2,66 s / 4 Hz	Les ours décidèrent d'aller se promener 2,17 s / 4 Hz	Les ours virent que leurs chaises avaient été utilisées 2,93 s / 6 Hz
Spanish	Los osos vivían juntos en una casa 2,60 s / 4 Hz	Los osos tomaron un gran paseo 2,19 s / 3 Hz	Los osos vieron sus sillas siendo utilizadas 3,00 s / 4 Hz

Table 2. Stimulus information. Three sets of sentences were used during the study. Each set contained one sentence translated into the three languages. The translations were slightly modified in order to match sentence duration across languages within the same set. The duration and syllabic rate is included for each sentence.

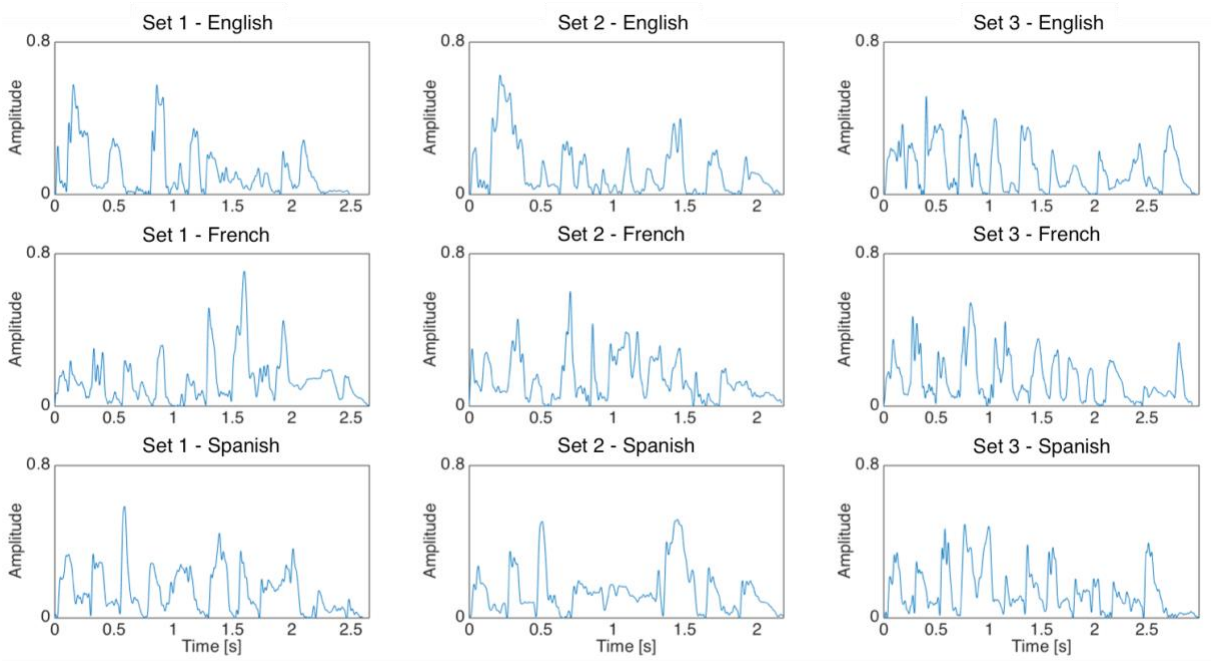


Figure 7. Speech envelopes of the nine sentences used in this study.

2.2.4 EEG Data acquisition

EEG data were recorded with active electrodes and an acquisition system from Brain Products (actiCAP & actiCHamp). A 10-channel layout was used to acquire cortical responses from the following scalp positions: F7, F3, FZ, F4, F8, T7, C3, CZ, C4, T8 (see Fig. 6b). The 10 channels were chosen to include those where auditory and speech perception related neural responses are typically observed in newborns (Stefanics et al., 2009; Tóth et al., 2017) (channels T7 and T8 used to be called T3 and T4, respectively). Additional electrodes were placed on both mastoids for online reference, and a ground electrode was placed on the forehead. Data were recorded at a sampling rate of 500 Hz.

2.2.5 Envelope Tracking Analysis

To investigate whether neural oscillations at birth already synchronize to the envelope of a speech signal during a passive listening task, we compared the amplitude and the phase of these two oscillatory signals. Synchronization to the amplitude and/or the phase of the speech envelope is a common measure for assessing speech envelope tracking in older populations (Abrams et al., 2008; Pefkou et al., 2017). Here, we computed both measures to explore which features of the speech envelope are encoded by the newborn brain. We pre-processed and analysed all the data using custom Matlab® scripts developed as part of this study. For the remaining of this document we will refer to amplitude synchronization and phase synchronizations as amplitude-encoding and phase-encoding respectively.

Amplitude-Encoding Analysis

When pre-processing the EEG and the speech signals, and analysing amplitude-encoding, we followed Abrams et al. (2008), and Kubanek et al. (2013). We started by band-pass filtering the raw EEG data between 1 and 40 Hz using a zero phase-shift Chebyshev filter, and then segmented it into a series of 2,960-ms long epochs. Each epoch started 200 ms before the utterance onset (corresponding to the pre-stimulus baseline), and contained a 2,760 ms long post-stimulus interval. Epochs were baseline corrected using the 200 ms pre-stimulus period before undergoing a three-stage process to reject the contaminated ones. First, epochs with amplitude exceeding ±75 μV were rejected. Second, if the standard deviation (SD) of an epoch was higher than 3 times the mean SD of all epochs, or lower than one-third the mean SD, the epoch was rejected. Third, the remaining epochs underwent visual inspection to remove any residual artifacts. After epoch rejection was completed, the evoked response was

computed by averaging all remaining epochs. This averaging was done per participant, per language and per channel. Each participant contributed in average with 49 epochs for the English condition, 52 for the French and 49 for the Spanish. Participants who had less than 20 remaining epochs after epoch rejection were not included in the data analysis (n=4, as also reported in the *Participants* section).

The amplitude envelopes of the speech signals were obtained by calculating the magnitude of the Hilbert transform of the stimulus waveforms. The envelopes were then low-pass filtered at 40 Hz and down-sampled to 500 Hz to match the characteristics of the evoked responses. The speech envelopes of the presented sentences are displayed in Fig. 7.

We assessed amplitude-encoding by performing cross-correlation analysis between the speech envelopes and the evoked responses at the participant level. The cross-correlation analysis was executed by computing the Spearman correlation between the two signals for lags between 80 ms and 600 ms (in 2 ms steps). At each step, the correlation was calculated for the "envelope-following" period (defined as the time range after 250 ms in Abrams et al., 2008) of both signals. The time range from 0 to 250ms was not included in the correlation as it was considered to be an onset period (as in Abrams et al., 2008). The latency of the peak of the cross-correlation function corresponds to the time lag at which the EEG response and the stimulus envelope are most similar, i.e. the time lag at which envelope tracking takes place. The cross-correlation analysis was performed per participant, per language and per channel. To ensure that an optimal lag could be found, we averaged the cross-correlation functions across the channels that were at least weakly correlated ($\rho>0.15$), as done by Kubanek et al. (2013), and obtained a mean time lag for each language at the participant level. Figure 8 presents examples of the average cross-correlation functions obtained for one subject for the three tested languages.

Figure 9 presents the time lags obtained per participant and per language. The peaks in the cross-correlation functions occur at similar latencies (time lags) for the three languages, as shown in Fig. 8 and Fig. 9. To confirm this statistically, we submitted the time lags from all the participants to a repeated measures ANOVA with language as a within-subjects factor. We did not find a main effect of language on the time lags [$F(2,92) = 0.044$, $p=0.957$], therefore we averaged the cross-correlation functions from the languages that were at least weakly correlated ($\rho>0.15$) to obtain a mean time lag for each participant (Fig. 8d).

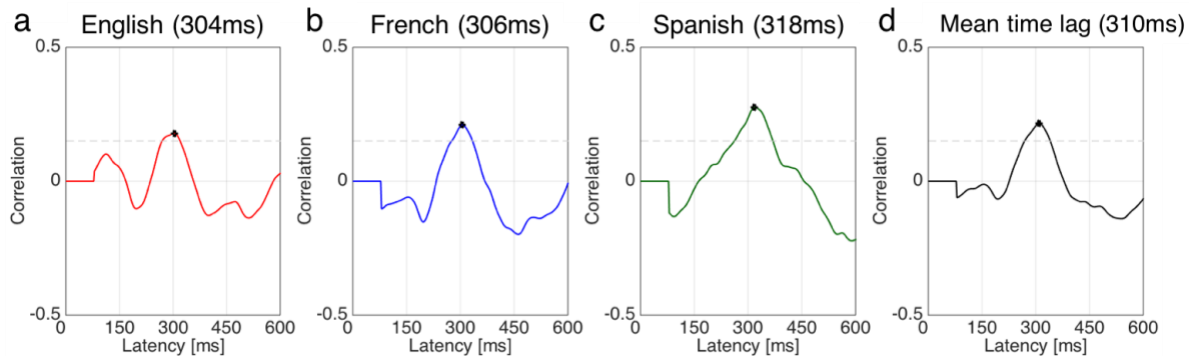


Figure 8. Average cross-correlation functions for one participant. **a)** Cross-correlation function for English; latency of the peak at 304 ms. **b)** Cross-correlation function for French; latency of the peak at 306 ms. **c)** Cross-correlation function for Spanish; latency of the peak at 318 ms. **d)** and the grand mean obtained by averaging across languages (black curve).

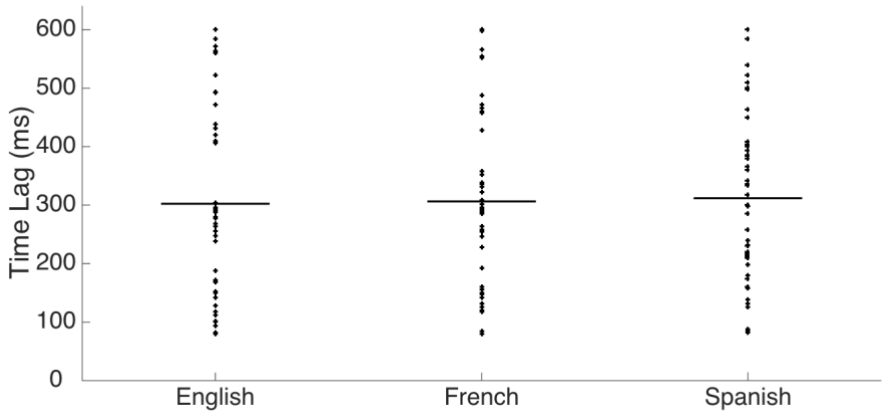


Figure 9. Time lags obtained at the participant level (per language). These values were submitted to a repeated measure ANOVA to test for the effect of language [F(2,92) = 0.044, p=0.957].

As a last step, we quantified amplitude-encoding as the correlation between the evoked responses (average of non-rejected trials) and the speech envelopes evaluated at the mean time lag obtained for each participant. It needs to be noted that evoked responses were extracted from the complete EEG signal (band-pass filtered, 1-40 Hz), and not from its envelope. The correlation values were Fisher transformed before statistical analysis as is customary when performing statistical comparisons of correlation coefficients. Even though only weakly correlated channels were included when obtaining the participants' time lags, all channels were included in the amplitude-encoding correlation analysis.

Statistical Analyses of Amplitude-Encoding Results

To validate our correlation results and establish envelope tracking, first we assessed if their distributions were significantly different from zero by submitting them to one-sample t-tests (two-tailed). This was done for each language and each channel. As a second step, we also compared our amplitude-encoding results against chance, by implementing a permutation test that shuffled the EEG data 1000 times along the time dimension while keeping the other labels intact (participant, language, and channel) (Hurtado et al., 2004; Sun et al., 2011). At each iteration, we applied the same procedure that was applied on the non-permuted data: 1) we ran a cross-correlation analysis between the permuted EEG and the speech envelopes, 2) we computed a time-lag, 3) we obtained the correlation values at the given time-lags, and 4) we Fisher transformed the correlation values. The results obtained from the permuted data served as a baseline of comparison for our amplitude-encoding results. The p-values for the permutation test were derived as the number of permutations out of the 1000 whose correlation distributions were not significantly different from the correlation distributions obtained for the non-permuted data (in paired samples t-tests). As a third statistical analysis we tested linear mixed effects models to assess if envelope tracking was different across languages and/or across channels. Table 3 presents the models that were run and compared in SPSS. The best fitting model was chosen based on the AIC (Akaike information criterion) value. The sample size for all the analyses previously described was 47 participants.

Model	Random Factors	Fixed Factors	Dependent Variable
m1	participant	language	correlation coefficients
m2	participant, language	language	correlation coefficients
m3	participant	channel	correlation coefficients
m4	participant, channel	channel	correlation coefficients
m5	participant	language, channel	correlation coefficients
m6	participant	language, channel, (language*channel)	correlation coefficients
m7	participant, language	language, channel	correlation coefficients
m8	participant, channel	language, channel	correlation coefficients

Table 3. List of tested linear mixed effects models.

Habituation Effect on Amplitude-Encoding Results

Even though investigating habituation effects is not an objective of this study, we assessed if the stimuli's repetitions introduced a habituation that could have had an effect on our correlation results. To do so, we averaged the first 5 non-rejected epochs of each language block, as well as the last 5 non-rejected epochs, for each participant separately. Then, we calculated the correlation between the speech envelope and each of these averaged EEG responses (EEG_beginning and EEG_end) by applying the mean time lag obtained previously for each participant in the *Amplitude-Encoding Analysis* section. The correlation values were Fisher transformed before statistical analysis. We submitted the correlation values from both conditions (EEG_beginning and EEG_end) to paired samples t-tests to evaluate the effect of habituation. This analysis was performed per language and per channel.

Phase-Encoding Analysis

To assess phase-encoding, we first determined the syllable rate of each sentence by detecting the number of peaks per second in their speech envelopes, and rounded the values to the closest integer (as performed by Pefkou et al., 2017). The syllabic rate of the tested sentences ranged from 3 to 6 Hz (see sentence details in Table 2). We band-pass filtered the raw EEG data from all participants between 3 and 6 Hz using a zero phase-shift Chebyshev filter. After filtering, the EEG data underwent the same processing steps as described in the *Amplitude-Encoding Analysis* section: epoching, baseline correction, epoch rejection, and averaging of remaining epochs to obtain the evoked response. The epochs rejected from the EEG data filtered at the syllabic rate were those identified as contaminated by the three-stage process run on the previous section (*Amplitude-Encoding Analysis*).

To match the characteristics of the filtered EEG, the speech envelopes were also band-pass filtered between 3 and 6 Hz. We quantified phase-encoding by computing the phase-coherence between the filtered EEG and the filtered speech envelopes by following these steps: 1) obtaining the Hilbert transform of both signals, 2) calculating the unitary signals of each analytic signal, and 3) computing the phase synchronization between both unitary signals. This analysis was performed per participant, per language and per channel.

Statistical Analyses of Phase-Encoding Results

To validate our phase-coherence results and establish phase-locking to the speech envelope, first we assessed if their distributions were significantly different from zero by submitting them to one-sample t-tests (two-tailed). This was done for each language and each channel. As a second step, we also compared our phase-encoding results against chance, by implementing a permutation test similar to the one used for the amplitude-encoding results. We permuted the EEG data 1000 times, and for each iteration we computed the phase-coherence between the permuted EEG and the speech envelopes. The phase-coherence results obtained from the permuted data served as a baseline of comparison against chance for the phase-encoding results. As a third statistical analysis we tested linear mixed effects models to assess if phase-locking was different across languages and/or across channels. Table 3 presents the models that were run and compared in SPSS. The best fitting model was chosen based on the AIC (Akaike information criterion) value. The sample size for all the analyses previously described was 47 participants.

2.2.6 Time-Frequency Analysis

To investigate the activity of the neural oscillations in the different tested conditions (familiar language, unfamiliar languages, resting state) we performed a time-frequency analysis. First we filtered the raw EEG signals with a 50-Hz notch filter to eliminate the power line noise. Then, we band-pass filtered the denoised EEG signals between 1 and 100 Hz to be able to measure all frequency bands of interest, and segmented them into a series of 2,564 ms long epochs. Each epoch started 400 ms before the utterance onset (corresponding to the pre-stimulus baseline), and contained a 2,164 ms long post-stimulus interval (corresponding to the duration of the shortest sentence). For the resting state data (3 minutes recorded during silence at the beginning of the experiment and 3 minutes at the end), we followed the same epoching procedure described above for the stimulation data, arbitrarily segmenting the 3 minutes into 2,564ms long epochs, which yielded 70 epochs for each rest condition. All the processing steps that follow were applied to the speech conditions (English, French, Spanish) and the resting state conditions (Rest_before, Rest_after) in a similar fashion. To reject contaminated epochs, we applied the same three-stage process implemented in the *amplitude-encoding analysis* section. Each participant contributed on average 49 epochs for the English condition, 52 for the French, 49 for the Spanish, 50 for Rest_before, and 56 for Rest_after. Participants who had less than 20 remaining epochs in a given condition after epoch rejection were excluded from the data analysis involving that condition. The 47 infants who entered final

data analysis were included in the language conditions (English, French, Spanish). However, some of these 47 infants did not provide a sufficient number of non-rejected trials in the Rest_before (n=4) and Rest_after (n=6) conditions, and were not included in some analyses.

The non-rejected epochs were subjected to time-frequency analysis to uncover stimulus-evoked oscillatory responses using a Matlab® toolbox named 'WTools' (provided to us upon request; Parise & Csibra, 2013). With this toolbox we performed a continuous wavelet transform of each non-rejected epoch using Morlet wavelets at 1 Hz intervals in the 1–100 Hz range. To remove the distortion introduced by the wavelet transform, the first and last 200 ms of each epoch were removed after the transformation. The transformed epochs were then baseline corrected using the average amplitude of the remaining 200 ms pre-stimulus window as baseline, subtracting it from the whole epoch at each frequency. Then, we removed the 200 ms baseline to obtain a 100 x 982 matrix of spectral amplitude values, corresponding to 100 frequency bins (from 1-100 Hz) and 982 time samples (1,964 ms, at a 500 Hz sampling frequency). This time-frequency analysis was performed for each non-rejected epoch, for each participant and for each channel. The time-frequency responses were then averaged across the non-rejected epochs and across the 10 channels for each condition separately to obtain the stimulus-evoked response. This yielded one time-frequency response matrix per condition for each participant. We then submitted the time-frequency response matrices to the following statistical analyses:

Analysis 1: Effect of Frequency Within a Condition

For this analysis we focused on the neural oscillations operating at 6 frequency bands of interest: delta, theta, alpha, beta, low-gamma, and high-gamma. Here we investigated how neural oscillations at these frequency bands behave within each language condition. To carry out this analysis we extracted from each time-frequency matrix the time series corresponding to the frequency bands of interest, and averaged them across the frequencies within each band: delta (1-3 Hz), theta (4-8 Hz), alpha (9-12 Hz), beta (13-30 Hz), low-gamma (31-50 Hz), high-gamma (51-80 Hz). Then, we averaged the mean time series of each band along the 982 time samples to obtain a mean spectral amplitude value for each frequency band. This analysis is thus similar to evaluating the spectrum of the EEG signal, specific to the 6 frequency bands of interest. To assess differences across frequency bands within each condition, we submitted the mean spectral amplitude values from all the participants to a repeated measures ANOVA with frequency band as the within-subjects factor. The sample

size was 47 participants for the language conditions (English, French, Spanish), 43 participants for the Rest_before condition, and 41 participants for the Rest_after condition.

Analysis 2: Effect of Language

Here we investigated whether neural oscillations behave differently across the tested languages. By assessing language differences, we aimed to explore which frequency bands (if any) reflect the rhythmic discrimination that takes place between the syllable-timed languages (French and Spanish) and the stress-timed language (English), as well as the familiarity that newborns have with the prenatally heard language (French). We explored language differences in the neural oscillations by submitting the spectral amplitude values of the 100 frequency bins (1-100 Hz) over the 982 time samples from the 3 languages to a repeated measures ANOVA with language as the within-subjects factor. As post hoc analyses, we computed paired samples t-tests (two-tailed) to compare the languages pairwise. The sample size was 47 participants.

Analysis 3: Language vs Resting State

Here we investigated whether neural oscillations already relate to speech processing at birth by assessing differences between neural oscillations during resting state and during speech processing. To do so, we compared the time-frequency data from each language condition to the resting state condition (Rest_before) using paired samples t-tests (two-tailed) as described in *Analysis 2*. This analysis was performed for each language condition separately. The sample size was 43 participants.

Analysis 4: Rest Before vs. Rest After

Here we investigated whether speech stimulation had a lasting effect on neural oscillations, or if neural oscillations go back to the resting state baseline after speech stimulation has taken place. To do so, we compared the time-frequency data from the two resting state conditions (Rest_before, recorded at the beginning of the experiment, and Rest_after, recorded at the end of the experiment) using paired samples t-tests (two-tailed) as described in *Analysis 2*. The sample size was 41 participants.

2.3 RESULTS

This section presents the results of the various analyses performed on the newborn EEG data.

2.3.1 Envelope Tracking

To assess speech envelope tracking we compared the EEG signals to the speech envelopes using two measures: amplitude-encoding and phase-encoding. For illustration, the average EEG time series from one participant are displayed in Figure 10.

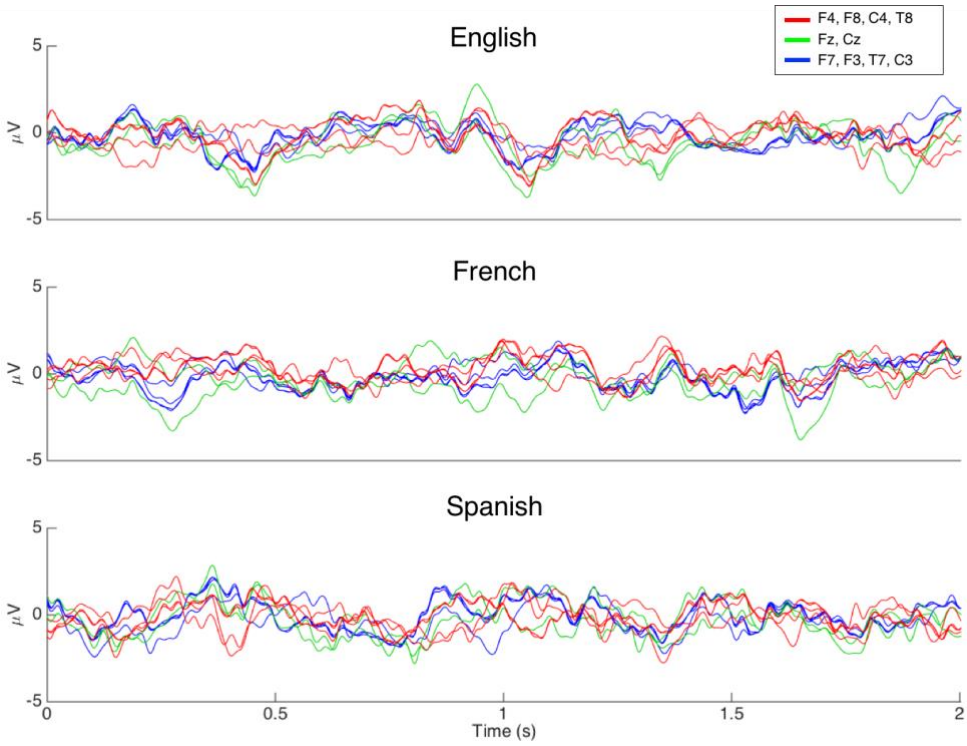


Figure 10. Average EEG time series from one newborn participant for the three language conditions. The red curves correspond to channels in the right hemisphere (F4, F8, C4, T8); the green curves to channels in the central region (Fz, Cz); and the blue curves to channels in the left hemisphere (F7, F3, T7, C3). These EEG responses correspond to the average of the non-rejected trials.

Amplitude-Encoding Results

Figure 11 presents the envelope tracking results of a single participant as illustration. In this figure we can observe how the EEG recordings (red, blue and green curves) track the speech envelopes (black curves) in the three languages.

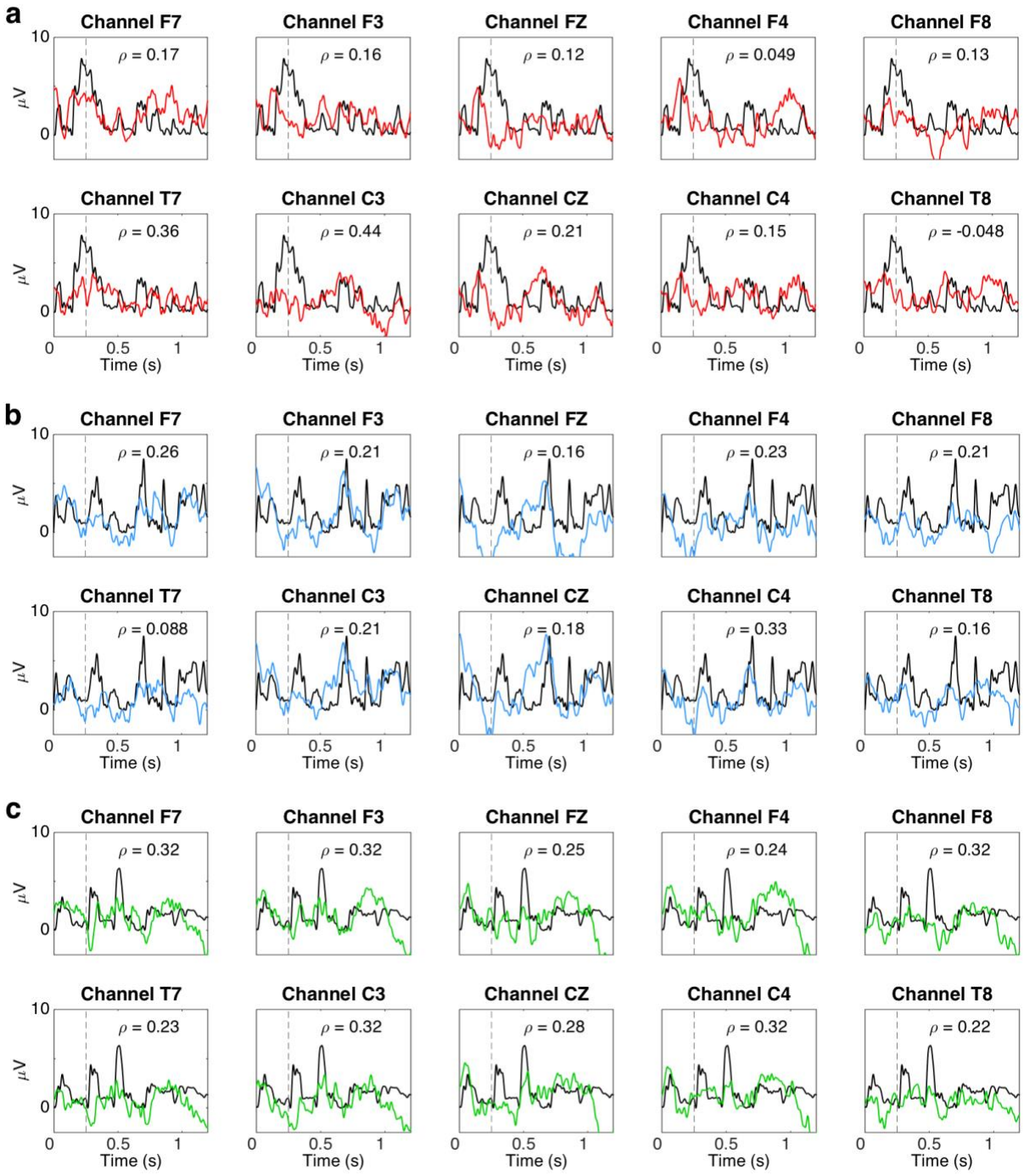


Figure 11. Amplitude envelope tracking of speech in an individual newborn participant. **a)** Envelope-tracking of an English sentence. The black curves represent the speech envelope and the red curves the cortical activity measured per channel. For visualization purposes the EEG data was shifted backward in time; this time shift corresponds to the time lag between the neural response and the speech signal. The Spearman correlation between the two signals is indicated by "ρ". **b)** Envelope-tracking of a French sentence. The blue curves represent the cortical activity measured per channel. **c)** Envelope-tracking of a Spanish sentence. The green curves represent the cortical activity measured per channel. The vertical line at 250 ms indicates the beginning of the envelope-following period.

Mean correlation values across the whole group of infants varied from channel to channel, ranging from 0.048 to 0.128 (mean 0.102; SD 0.0248) for **English**, from 0.090 to 0.142 (mean 0.126; SD 0.0195) for **French**, and from 0.059 to 0.152 (mean 0.122; SD 0.0285) for **Spanish**. These group results are listed in Table 4 and displayed in Fig. 12 (yellow boxes) per channel and per language. The correlation distributions were significantly higher than zero for most channels and languages when submitted to one-sample t-tests (two-tailed). The p-values from these t-tests were corrected for multiple comparisons (i.e. to adjust the family-wise error rate for each comparison) using the Bonferroni correction, and their details are listed in Table 4. Only channel T8 was not significantly different from zero for the English and Spanish conditions.

	ENGLISH			FRENCH			SPANISH		
Channel	mean	p-value	t-value	mean	p-value	t-value	mean	p-value	t-value
F7	0.121	0.00	5.118	0.121	0.00	4.627	0.120	0.00	4.453
F3	0.116	0.00	5.520	0.123	0.00	5.025	0.147	0.00	6.096
Fz	0.097	0.00	4.374	0.124	0.00	5.178	0.152	0.00	6.297
F4	0.128	0.00	5.551	0.140	0.00	5.605	0.142	0.00	5.189
F8	0.112	0.00	4.825	0.142	0.00	5.847	0.120	0.00	4.287
T7	0.075	0.05	2.921	0.095	0.00	3.771	0.088	0.01	3.705
C3	0.111	0.00	4.265	0.141	0.00	6.331	0.124	0.00	5.389
Cz	0.120	0.00	5.451	0.140	0.00	5.891	0.129	0.00	5.015
C4	0.091	0.00	4.134	0.141	0.00	6.256	0.135	0.00	5.086
T8	0.048	0.54	1.981	0.090	0.00	3.764	0.059	0.09	2.717

Table 4. Group mean correlation values for the newborn participants per channel and per language. The correlation distributions were submitted to one-sample t-tests (two-tailed) to compare them to zero. The p-values (Bonferroni corrected) and t-values of these t-tests are presented here per channel. Non-significant results are highlighted in red.

To further establish envelope tracking, we used a permutation test to compare against chance. We permuted the EEG data 1000 times, and performed the envelope tracking analysis at each permutation. The results from the permutation test are displayed in Fig. 12 (blue boxes) next to the results from the measured data (yellow boxes). As it can be seen in Fig. 12, most correlation distributions from the measured data are significantly higher than the distributions from the permuted data for the three languages, with two exceptions: channels T7 and T8 do not perform envelope tracking better than chance in at least one language (T7 for French, and T8 for the three languages). These results thus establish that the newborn

brain encodes the amplitude of the speech envelope in all three languages more accurately than what would be expected purely by chance.

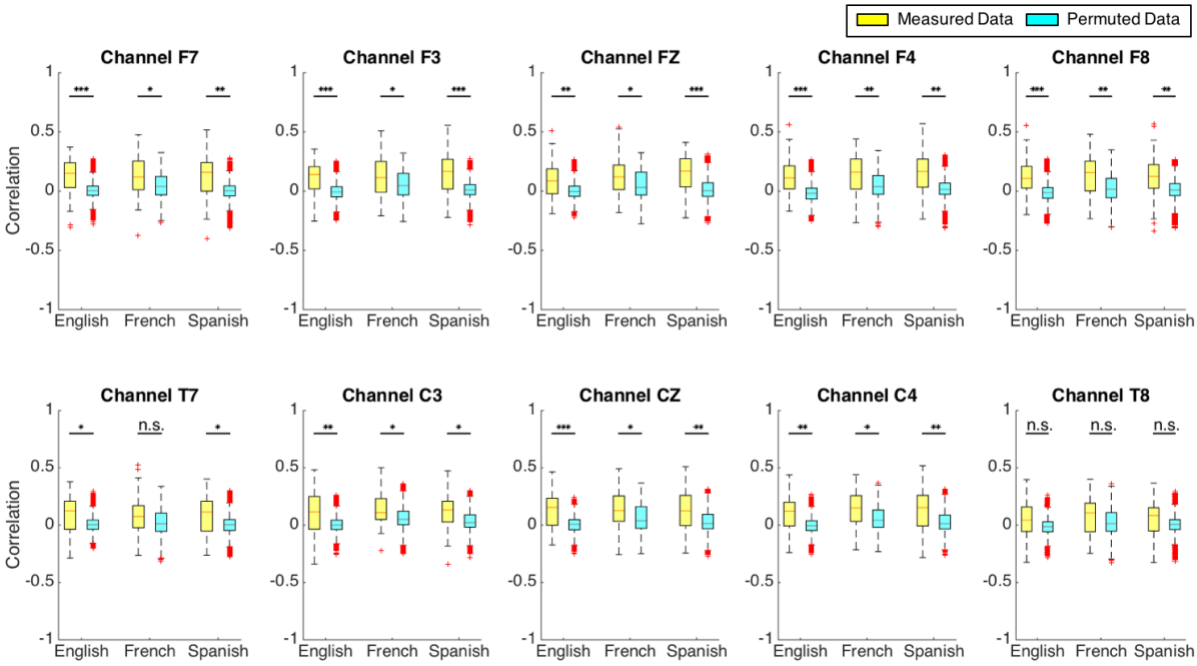


Figure 12. Amplitude envelope tracking results for the newborn participants are displayed for each channel, and for the three languages in a box and whisker plot. The yellow boxes represent the distributions of the results from the measured EEG data, and the blue boxes represent the distributions of the results from the permuted data. n.s.: non-significant, * p<0.05, ** p<0.01, *** p<0.001.

To assess possible differences in the amplitude-encoding results across languages and channels, we implemented linear mixed effects models with the correlation values as the dependent variable (see Methods). The best fitting model (model '*m2'*, Table 3) had *Language* as fixed factor, and *Participant* and *Language* as random factors. For this model, there was no main effect from the fixed factor *Language* [F(2,140) = 0.645, p = 0.526].

Habituation Effect on Amplitude-Encoding Results

To evaluate if habituation to the sentence due to the repetitions took place and had an effect on our amplitude-encoding results, we computed envelope tracking for trials from the beginning and for trials from the end of each language block. Figure 13 displays the group results per channel, per language, and per condition (EEG_beginning and EEG_end). Results from the first 5 non-rejected trials are represented by the green boxes, while results from the last 5 non-rejected trials are represented by the purple boxes. As it can be seen in Fig. 13,

most channels show no significant difference between the EEG responses at the beginning and at the end of the language block. **Spanish** is the only language that shows an effect due to the repetitions (channels F8, C4, T8) having a higher correlation towards the end of the language blocks. For the case of channel T8, we found that amplitude-encoding results were not significantly higher than chance (Fig. 12), so even if the repetitions boosted the correlation responses, their effect was not strong enough for envelope tracking to take place. In the absence of a considerable habituation effect, we did not pursue the habituation issue further in the subsequent analyses.

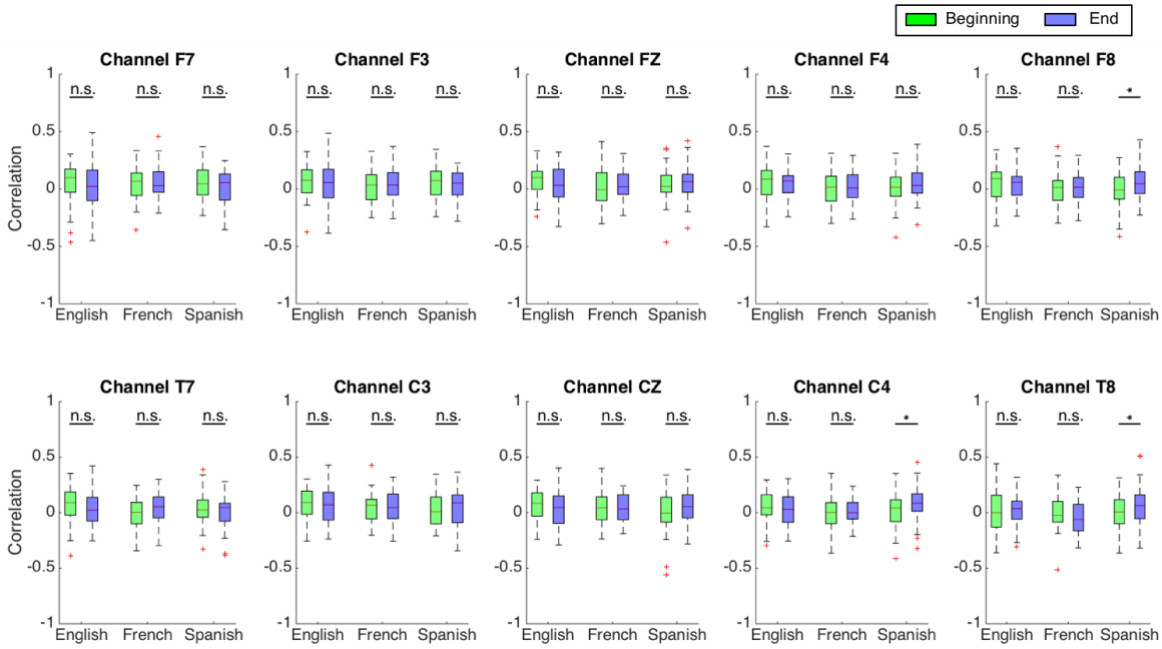


Figure 13. Amplitude envelope tracking results at the participant level for trials taken from the beginning of the language block and for trials taken from the end of the language block. The green boxes represent the distribution of the results from the first 5 non-rejected trials, and the purple boxes represent the distribution of the results from the last 5 non-rejected trials. Results are displayed for each channel, and for the three languages. n.s.: non-significant, * $p<0.05$.

Phase-Encoding Results

Mean phase-coherence values across the whole group of infants varied from channel to channel, ranging from 0.292 to 0.350 (mean 0.319; SD 0.0200) for **English**, from 0.302 to 0.349 (mean 0.331; SD 0.0147) for **French**, and from 0.312 to 0.377 (mean 0.343; SD 0.0198) for **Spanish**. These group results are listed in Table 5 and displayed in Fig. 14

(yellow boxes) per channel and per language. The phase-coherence distributions were significantly higher than zero for all channels and languages, when submitted to one-sample t-tests (two-tailed). The p-values from these t-tests were corrected for multiple comparisons (i.e. to adjust the family-wise error rate for each comparison) using the Bonferroni correction, and their details are listed in Table 5.

	ENGLISH			FRENCH			SPANISH		
Channel	Mean	p-value	t-value	mean	p-value	t-value	mean	p-value	t-value
F7	0.350	0.00	15.263	0.338	0.00	15.591	0.355	0.00	14.622
F3	0.298	0.00	13.781	0.339	0.00	15.016	0.343	0.00	14.648
Fz	0.321	0.00	13.544	0.335	0.00	14.534	0.344	0.00	15.508
F4	0.338	0.00	15.368	0.338	0.00	14.322	0.312	0.00	12.450
F8	0.334	0.00	12.986	0.337	0.00	14.769	0.332	0.00	13.020
T7	0.314	0.00	13.917	0.332	0.00	13.168	0.377	0.00	14.199
C3	0.304	0.00	12.367	0.349	0.00	13.964	0.368	0.00	16.661
Cz	0.302	0.00	12.700	0.302	0.00	12.169	0.342	0.00	15.117
C4	0.337	0.00	13.274	0.309	0.00	13.979	0.321	0.00	13.106
T8	0.292	0.00	12.120	0.327	0.00	14.882	0.337	0.00	14.188

Table 5. Group mean phase-coherence values for the newborn participants per channel and per language. The phase-coherence distributions were submitted to one-sample t-tests (two-tailed) to compare them to zero. The p-values (Bonferroni corrected) and t-values of these t-tests are presented here per channel. All channels were significantly higher than zero for the three languages.

To further establish phase-encoding we used a permutation test, similar to the one implemented for the amplitude-encoding results, to compare against chance. We permuted the EEG data 1000 times, and performed the phase-encoding analysis at each permutation. The results from the permutation test are displayed in Fig. 14 (blue boxes) next to the results from the measured data (yellow boxes). As it can be seen in Fig. 14, our phase-coherence results are significantly higher than the permutation results for all the channels and all the languages. These results thus establish that the newborn brain encodes the phase of the speech envelope in all three languages more accurately than what would be expected purely by chance.

To assess possible differences in the phase-encoding results across languages and channels, we implemented linear mixed effects models with the phase-coherence values as the dependent variable (see Methods). The best fitting model (model '*m2'*, Table 3) had

Language as fixed factor, and *Participant* and *Language* as random factors. For this model, there was no main effect from the fixed factor *Language* [F(2,163) = 1.388, p = 0.252].

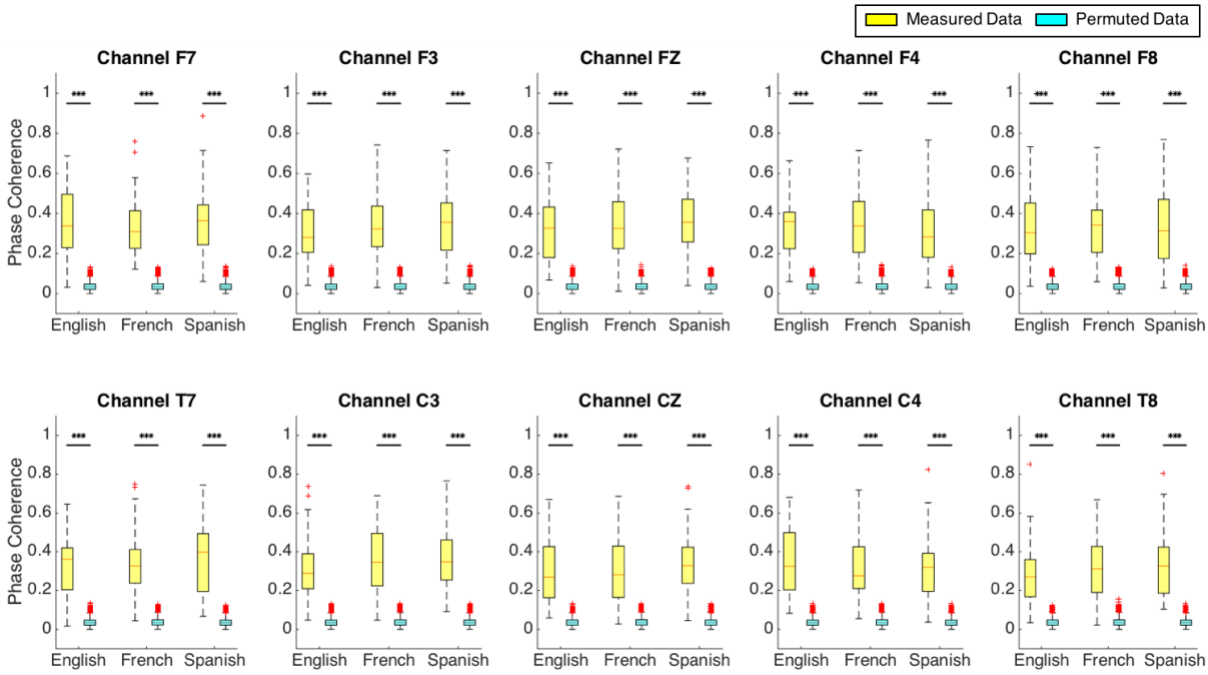


Figure 14. Phase-coherence results for newborn participants are displayed for each channel, and for the three languages in a box and whisker plot. The yellow boxes represent the distributions of the results from the measured EEG data, and the blue boxes represent the distributions of the results from the permuted data. *** p<0.001.

2.3.2 Time-Frequency Analysis

In order to investigate the behaviour of neural oscillations in the newborn brain, we obtained the time-frequency responses for the five tested conditions (English, French, Spanish, Rest_before, Rest_after), and submitted them to various statistical analyses to explore differences across frequencies and languages, as well as along time. Group mean time-frequency responses are displayed in Figure 15. In this section we will be plotting the spectral amplitude of the time-frequency responses (the square root of the power).

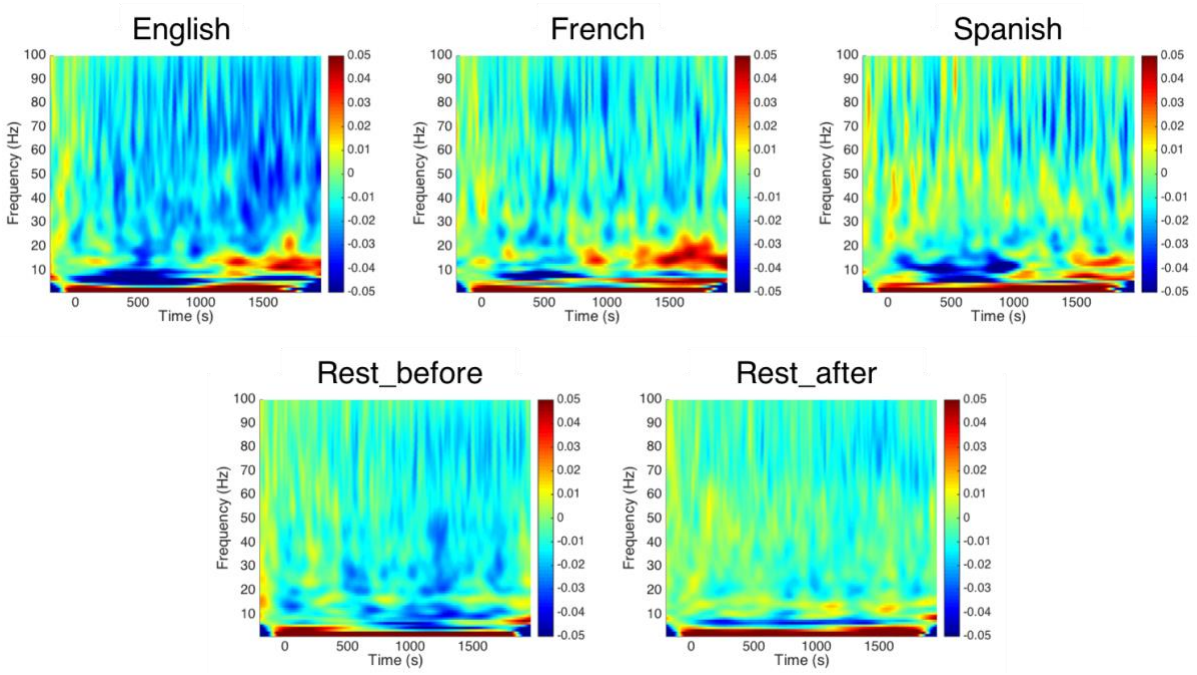


Figure 15. Group mean time-frequency responses for the five tested conditions: English, French, Spanish, Rest_before (resting state at the beginning of the experiment), and Rest_after (resting state at the end of the experiment). Each plot is the average of the non-rejected epochs, the 10 channels and the analysed participants for each condition: 47 infants for English, French and Spanish; 43 infants for Rest_before; 41 infants for Rest_after. The color bar on the right of each plot shows the spectral amplitude scale.

Analysis 1: Effect of Frequency Within a Condition

Figure 16 displays the mean spectral amplitude values per frequency band for each condition. To investigate if neural oscillations operating at the frequency bands of interest behave differently within each condition, we applied a repeated measures ANOVA with frequency band as a within-subjects factor. For every condition (tested separately), we found a main effect of frequency band on the mean spectral amplitude values. This effect was due to the **delta band** being significantly higher than all the other bands: for **English** [$F(2.448,112.620) = 33.975$, $p=0$], for **French** [$F(2.710,124.667) = 43.766$, $p=0$], for **Spanish** [$F(2.516,115.733) = 37.450$, $p=0$], for **Rest_before** [$F(2.622,110.138) = 150.008$, $p=0$], and for **Rest_after** [$F(2.256,90.246) = 133.056$, $p=0$]. These results were not corrected for multiple comparisons (one ANOVA per condition).

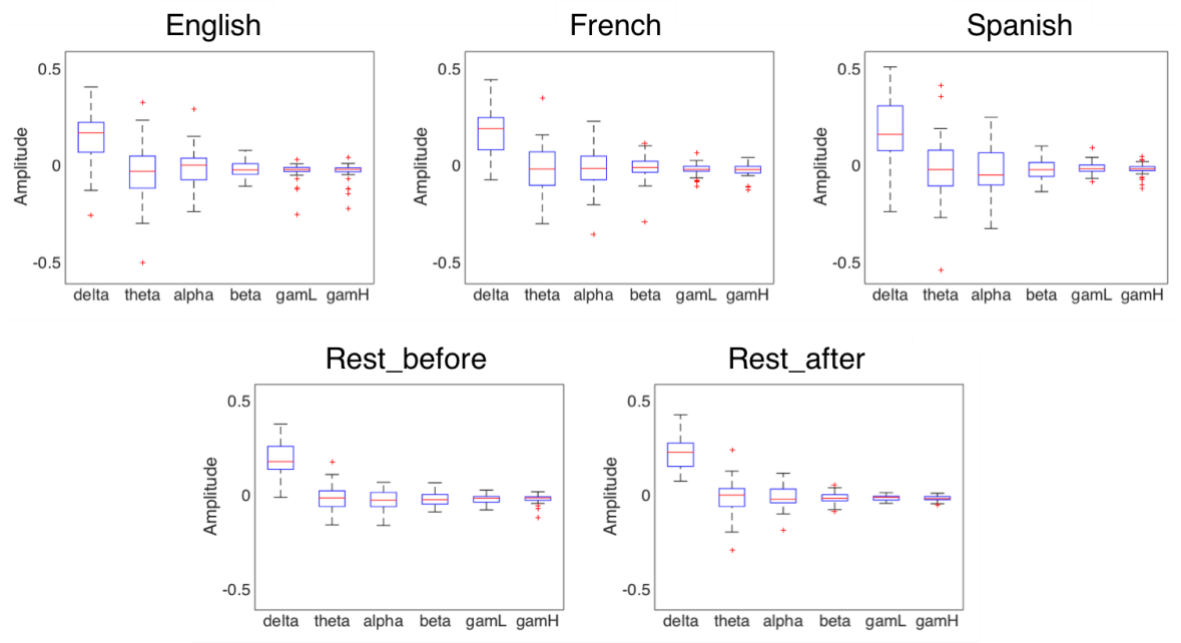


Figure 16. Mean spectral amplitude values for the six frequency bands of interest (delta, theta, alpha, beta, low-gamma, high-gamma) per condition (English, French, Spanish, Rest_before, Rest_after). For every condition, the delta band was significantly higher than all the other bands (Analysis 1).

Analysis 2: Effect of Language

In this analysis we assessed how neural oscillations differ across the three tested language conditions (English, French, Spanish) over the time course of stimulation. Figure 17a displays a time-frequency matrix with the results from the one-way repeated measures ANOVA that had language as the within-subjects factor (uncorrected for multiple comparisons). The red regions represent the time-frequency windows where the amplitude of the neural oscillations was significantly different across the 3 languages. Language differences were found in theta, alpha, beta, low-gamma and high-gamma bands, with some bands displaying a stronger effect than others: the language effect is strongest in the theta and low-gamma bands.

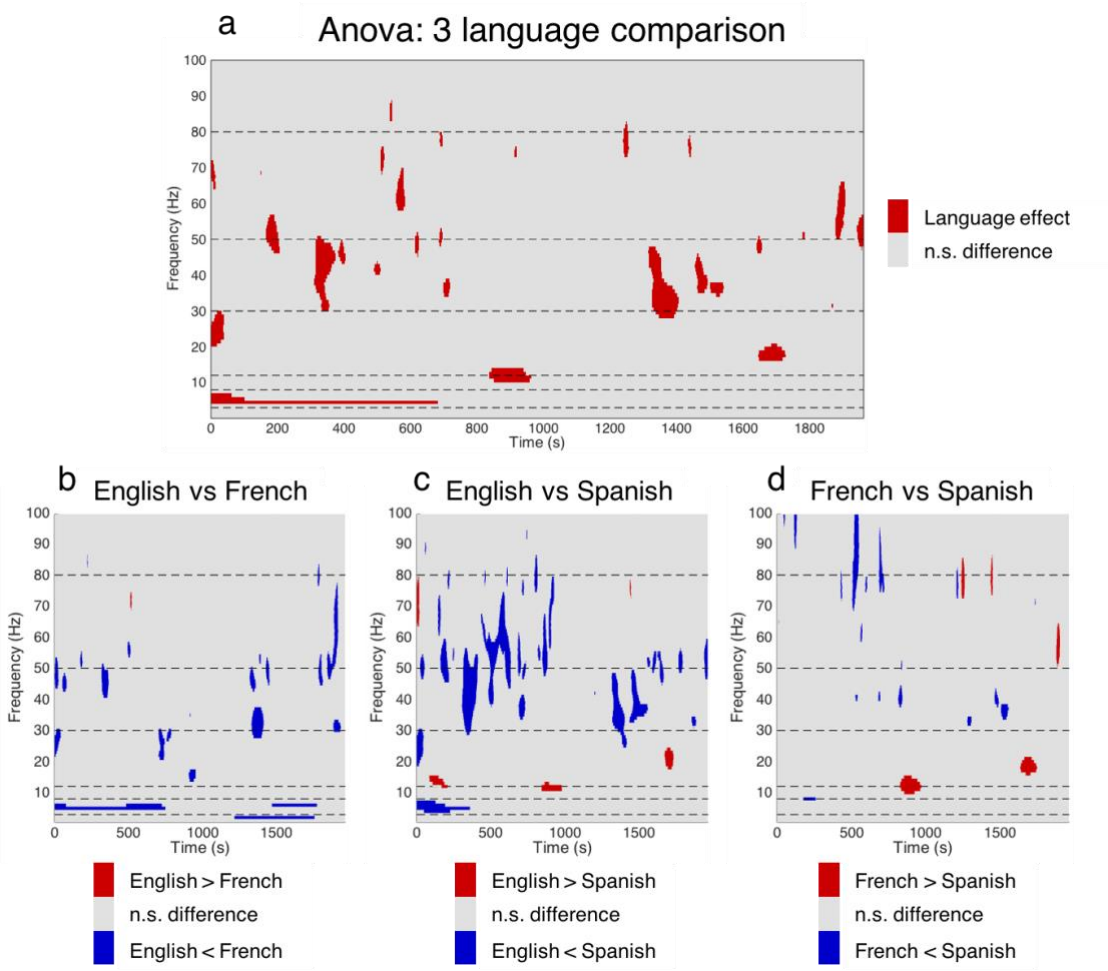


Figure 17. Time-frequency responses of significant language effect. **a)** Results of an ANOVA test between the neural responses of the 3 languages. The red regions represent the time-frequency windows where the amplitude of the neural oscillations was significantly different across the 3 languages. **b)** Results of a t-test between the neural responses of English and French. c) Results of a t-test between the neural responses of English and Spanish. D) Results of a t-test between the neural responses of French and Spanish. The coloured regions represent the time-frequency windows where the amplitude of the neural oscillations was significantly different across the 2 languages (Analysis 2).

To further explore the origin of these language differences we compared each pair of languages applying paired samples t-tests. Figure 17 (b,c,d) displays the time-frequency differences for each pairwise comparison. From these comparisons we can point out the following observations:

- Differences in the **theta** band (4-8 Hz) are only present in the comparisons between rhythmically different languages (English vs. French, and English vs. Spanish), and during the first 800 ms after stimulus onset.

- Only very small differences are present in the **alpha** band (9-12 Hz).
- Differences in the **beta** band (13-30 Hz) are present in all the comparison (in restricted time windows).
- Differences in the **low-gamma** band (31-50 Hz) are mostly present in the comparisons between rhythmically different languages (English vs. French, and English vs. Spanish).
- Differences in the **high-gamma** band (51-80 Hz) are mostly present in the comparison between English and Spanish.

Analysis 3: Language vs. Resting State

In this analysis we assessed how neural oscillations during speech processing differ from resting state neural oscillations at birth. Figure 18 displays the time-frequency differences between each language condition and the resting state condition obtained using paired samples t-tests (two-tailed, uncorrected for multiple comparisons). From these comparisons we can point out the following observations:

- The neural oscillations elicited by English and French in the **delta** band (1-3 Hz) have lower amplitude than those generated during resting state.
- In the **theta** band (4-8 Hz), oscillations elicited by English are significantly lower than those from the resting state, while the oscillations elicited by French are significantly higher than those from the resting state.
- In the **alpha** band (9-12 Hz), oscillations elicited by Spanish are lower than at rest (in a time window), while for English they are higher than at rest.
- In the **beta** band (13-30 Hz), oscillations elicited by English and French are higher than at rest, while for Spanish they are lower.
- In the **low-gamma** band (31-50 Hz), oscillations for English are lower than at rest, while for Spanish they are higher than for rest. For French, oscillations are higher than at rest at the onset (same as for Spanish).
- In the **high-gamma** band (51-80 Hz), the biggest differences are found for English, having lower amplitude oscillations than at rest.

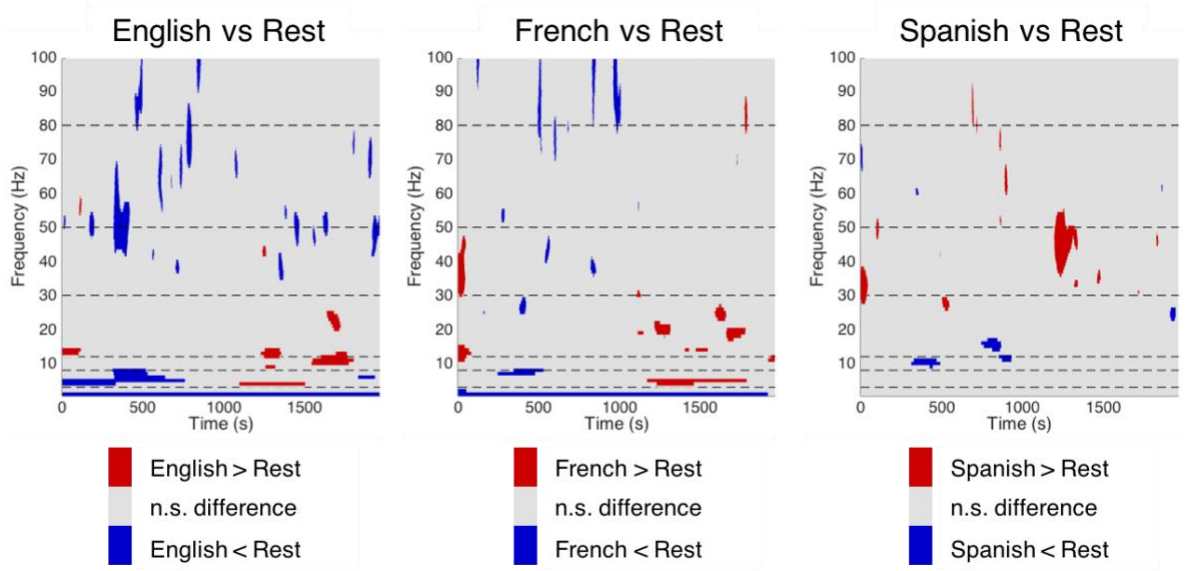


Figure 18. Time-frequency responses of language comparisons against resting state. Each matrix plot displays the time-frequency windows where the neural oscillations elicited by each language are significantly different from the neural oscillations at resting state. The red regions represent higher amplitude oscillations for the language condition (than rest), and the blue regions represent lower amplitude oscillations for the language condition (Analysis 3).

Analysis 4: Rest Before vs. Rest After

In this analysis we investigated if neural oscillations go back to the resting state baseline after speech stimulation takes place. Figure 19 displays the time-frequency differences between the two resting state conditions: Rest_before and Rest_after. From this comparison it can be observed that neural oscillations during the final resting state recording have higher amplitudes than during the initial resting state recording. Even though these differences occur during small time-frequency windows, they are present in all the frequency bands of interest (delta, theta, alpha, beta, low-gamma bands, high-gamma band).

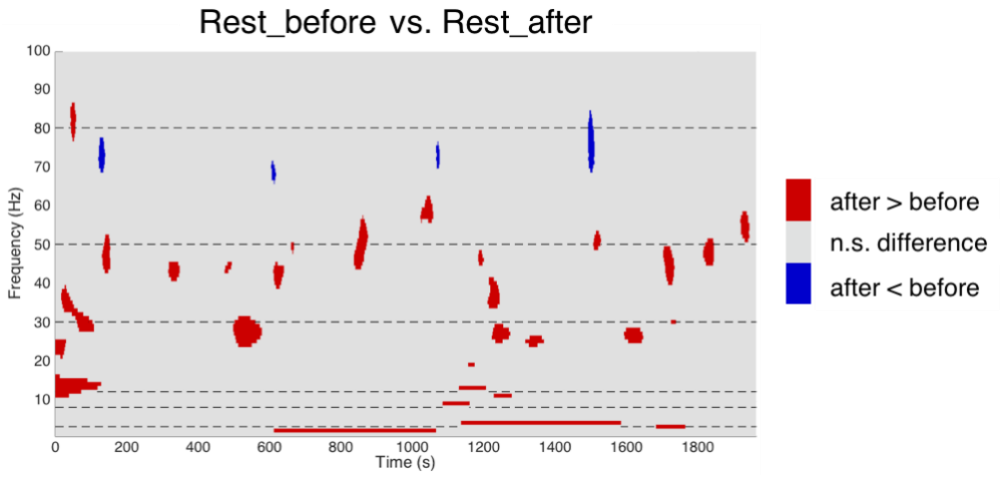


Figure 19. Time-frequency response of comparing the two resting state conditions: Rest_before vs. Rest_after. The red regions represent higher amplitude oscillations for the final resting state conditions, and the blue regions represent lower amplitude oscillations for the final resting state condition (Analysis 4).

Summary of Time-Frequency Results

Table 6 summarizes the results from all the time-frequency analyses presented in this chapter.

Analysis	Type of data	Effect Tested	Results
1	6 frequency bands	Frequency (per condition)	delta > other bands
2	full-band time-frequency data	Language	the language effect is strongest in the theta and low-gamma bands.
3	full-band time-frequency data	Language vs. Rest	delta: English, French < rest theta: English < rest French > rest alpha: English > rest Spanish < rest beta: English, French > rest Spanish < rest low-gamma: English < rest Spanish, French > rest high-gamma: English < rest
4	full-band time-frequency data	Rest_before vs. Rest_after	Rest_after > Rest_before (all bands)

Table 6. Summary of results for the time-frequency analyses.

2.4 DISCUSSION

In this chapter we investigated neural oscillations at birth in 5 different conditions: (1) during resting state, and prior to any auditory stimulation, (2) while listening to sentences in their native language, (3) in a rhythmically similar unfamiliar language, (4) in a rhythmically different unfamiliar language, and (5) again at rest, after the auditory stimulation had taken place. By exploring how the neural oscillations behave in the different conditions we aimed to investigate: (i) whether neural oscillations are in place at birth, and if they are, (ii) how do they relate to speech processing in the newborn brain, (iii) whether familiarity with the prenatally heard language has an effect on their behaviour, (iv) whether neural oscillations relate to the discrimination of rhythmically different languages, (v) whether speech envelope tracking is present at birth, and if it is, (vi) whether it is shaped by the prenatal experience. To tackle these questions, we analysed the newborn EEG data using various methodologies: speech envelope tracking by means of amplitude-encoding and phase-encoding, as well as time-frequency analysis investigating neural oscillations. Here, we will start by discussing the results of each analysis separately, and then we will bring them together to answer the proposed questions.

Amplitude-Encoding Analysis

To assess amplitude-encoding, we computed the correlation between the speech envelopes and the EEG evoked responses in the 1-40 Hz range. Our mean correlation results ranged from 0.048 to 0.128 for **English**, from 0.090 to 0.142 for **French**, and from 0.059 to 0.152 for **Spanish**. We validated our results by verifying if their distributions were significantly different from zero. We found that only channel T8 was not significantly different from zero for English and Spanish. We also submitted the correlation coefficients to a permutation test because we consider that it represents a more stringent way to evaluate the significance of our results. With this test we found that results from only two channels were not significantly higher than the permutation baseline: T7 for French, and T8 for the three languages. The permutation results are fairly consistent with the t-test results. Based on the fact that correlation distributions from most channels were significantly higher than zero (9/10) and significantly higher than chance (8/10), as well as the fact that we did not find a language effect on our results, we conclude that the newborn brain is able to track the amplitude of the speech envelope of familiar and unfamiliar languages equally well.

In the literature, not many studies have evaluated envelope tracking by means of amplitude-encoding, and to our knowledge, there are no infant studies evaluating this phenomenon. Therefore, we will relate our results to two studies that even though differ methodologically from our own in some ways, served us as reference when designing our experiment and analysing our data. When older children (9 to 13 years old) were tested in a similar paradigm (Abrams et al., 2008), a mean correlation of 0.37 was found. We believe there are two main reasons that explain why our results are lower than those found by Abrams et al, (2008). First, EEG recordings are less structured in newborns than in older populations: due to lower synaptic efficiency, young infants' neural waveforms are slow and do not show many well defined peaks (de Haan, 2007; Britton et al., 2016). It is the presence of peaks that facilitates correlating the EEG recordings to the speech envelopes, therefore the fact that newborn EEG does not exhibit well defined peaks, represents a challenge for this measurement. The second reason that explains why our results are lower is that the study with older children presented 1,000 repetitions of one sentence (compared to 100 repetitions in our study). By averaging a greater number of epochs, the resulting evoked response has a higher signal to noise ratio, which enhances the features that correlate to the speech envelope. Therefore, even though our results are lower than those found in the above mentioned study, it is not unexpected given our population. On a different note, our correlation results are of similar magnitude to those observed in adults tested with ECoG (electrocorticography, Kubanek et al., 2013) where mean correlation values ranged from 0.05 to 0.20. Thanks to the fact that ECoG is a more sensitive technique than scalp EEG, they were able to obtain these results during single stimulus presentations, compared to our study, where we used repetitions.

Habituation Effect on Amplitude-Encoding

In order to increase the signal to noise ratio in the neural responses, we presented sentences in repetitions in the same way as done in the study with older children (Abrams et al., 2008). To assess if these repetitions had an effect on our correlation results, we compared envelope tracking from 5 non-rejected trials from the beginning of each language block to envelope tracking from 5 non-rejected trials from the end of the same block. If habituation took place, we would expect to see lower envelope tracking towards the end of the language blocks, as the overall amplitude of the responses would decrease, making the signal less able to match the auditory envelope. We only found significant differences in 3 channels for a single

language, and they were in the opposite direction: envelope tracking (correlation) results were higher for the trials taken from the end of the block than from the trials taken from the beginning. If anything, we could argue that the repetitions might have helped improve envelope tracking. However, only 2 of the 3 channels that showed the effect was significant in the permutation test; therefore, even if repetitions boosted our correlation results, their effect was not strong enough to make a significant difference.

It needs to be noted that data rejection could impact which trials go into the analysis, such that the first 5 non-rejected trials at the beginning of a language block are not necessarily the first 5 recorded trials; they could have been taken from later in the block, once habituation could have already taken place. Considering the challenging task of recording infant EEG, this is a limitation we could have not avoided. This study was not designed to investigate habituation effects, and therefore this paradigm does not really allow us to determine if an early habituation effect took place. Nevertheless, our analysis suggests that even if there was one, it does not impact our envelope tracking results.

Phase-Encoding Analysis

To assess phase-encoding, we computed the phase-coherence between the speech envelopes and the EEG evoked responses in the 3-6 Hz range (syllabic rate). Our mean phase-coherence results ranged from 0.292 to 0.350 for **English**, from 0.302 to 0.349 for **French**, and from 0.312 to 0.377 for **Spanish**. We validated our results by verifying if their distributions were significantly different from zero. We found that results from all channels were significantly higher than zero for the 3 languages. We also submitted the phase-coherence values to a permutation test because we consider that it represents a more stringent way to evaluate the significance of our results. With this test we found that results from all channels were significantly higher than the permutation baseline for the 3 languages. The permutation results are consistent with the t-test results. Based on the fact that phase-coherence results were significantly higher than zero and significantly higher than chance, as well as the fact that we did not find a language effect on them, we conclude that the newborn brain is able to track the phase of the speech envelope of familiar and unfamiliar languages equally well.

To our knowledge, there are no infant studies evaluating envelope tracking by means of phase-encoding. Therefore, we will compare our results to one of the few adult studies relating the oscillatory phase directly to the phase of the speech envelope, as we did here. Our phase-coherence results are of similar magnitude to those observed in adults who were

presented with utterances in their native language (mean phase-coherence of around 0.27, Pefkou et al., 2017). Furthermore, in the same study, participants were also presented with time-compressed speech, exhibiting similar phase-encoding responses for intelligible and unintelligible speech. Our results for familiar and unfamiliar languages are then in line with those findings.

Time-Frequency Analysis

To investigate neural oscillations and their relationship with speech processing, we analysed the EEG signal in the 1-100 Hz range. We transformed the EEG responses into time-frequency representations (using Morlet wavelets) to obtain speech evoked neural oscillations, as well as resting state neural oscillations. Table 6 summarizes the results obtained from the various analyses that we conducted on our time-frequency responses.

When investigating the behaviour of the 6 frequency bands of interest (delta, theta, alpha, beta, low-gamma, and high-gamma) within each tested condition (analysis 1), we found that the amplitude of the neural oscillations in the delta band was higher than in all the other frequency bands. This result is expected, as brain activity measured with different recording systems (EEG, MEG, fMRI) exhibits a 1/f power spectrum, with slow frequencies having stronger power than high frequencies (He, 2014).

When evaluating the effect of language familiarity on the neural oscillations (analysis 2), we found the strongest effect in the theta and low-gamma bands (Fig. 17a). To further explore the origin of these differences, we compared the time-frequency responses of each pair of languages, and found that the theta and low-gamma effects are due to the differences between the rhythmically different languages (English vs. French, and English vs. Spanish). Neural oscillations in the theta band are significantly higher for the rhythmically familiar languages (French and Spanish) than for the rhythmically unfamiliar one (English), and exhibit no significant differences between the rhythmically similar languages (French and Spanish). These results likely reflect two things: (1) the preference newborns have for the prenatally heard language, and (2) the ability they have to discriminate rhythmically different languages, but not rhythmically similar ones (Mehler et al., 1988; Moon et al., 1993; Nazzi et al., 1998; Ramus et al., 2000). Prenatal experience is prosodic, influencing mostly the lower frequencies (delta and theta), therefore the greater theta activation found for French (and Spanish) is most likely a result of familiarity, i.e. prenatal experience. Moreover, the fact that theta activation is significantly different between the syllable-timed languages (French and Spanish) and the

stress-timed language (English) most likely reflects rhythmic discrimination. We want to point out the timing of the cross-linguistic effects on the theta band: the rhythmic differences are present early after sentence onset, which could be an indication that rhythm is encoded from the start, and then the system enters a specific rhythmic mode (e.g. syllable-timed), so it is no longer tracked once it is established, and probably once the other higher frequencies (specifically gamma) entrain to it. Considering that rhythm correlates with syllabic rate (Varnet et al., 2017), and syllabic rate relates to the theta band (Giraud & Poeppel, 2012), it makes sense to find rhythmic discrimination being reflected in the theta oscillations. Here we also found some differences in the low-gamma band for the comparison of rhythmically different languages, likely reflecting rhythmic discrimination. This is in line with what has been found in 3 month-olds (Peña et. al., 2010), where this effect was also observed in the gamma band but in somewhat higher frequencies, possibly due to differences in the experimental design (more sentences), the details of the time-frequency analysis, or developmental differences between the tested populations (newborns vs. 3-month-olds).

When assessing the effect of speech stimulation on the resting state neural oscillations (analysis 3) we observed different results for the three tested languages. For English, we found lower overall activation across most of the spectrum (delta, theta, low-gamma, high-gamma) compared to rest. For Spanish, we observed the opposite behaviour, having higher overall activation than rest (mostly in low-gamma). For French, differences are less marked, being both higher and lower at different time-frequency intervals. At the moment we do not understand these differences, but we suggest that it would be interesting to further investigate if language familiarity and the acoustic properties of the sentences could help explain these results. The fact that we found lower activation for English (the rhythmically unfamiliar language) is a promising indication that language familiarity could be involved. We will further look into each set of sentences separately to disentangle these effects.

When investigating the after-effects of speech stimulation (analysis 4) we explored whether neural oscillations would go back to the resting state baseline immediately after the auditory stimulation concluded. We observed that neural activity in the final resting state was higher in amplitude than the initial one. This finding is in line with what was found for 3-month-olds tested with NIRS (near-infrared spectroscopy, Homae et al., 2011), who exhibited higher connectivity in the post-stimulus resting period compared to the pre-stimulus one. A possible explanation to this phenomenon could be the occurrence of some short-term learning that is taking place while the neural oscillations continue to be active. This could be an

indication that the neural mechanisms supporting learning and neural plasticity are already operational at birth.

Research Questions & Answers

We can now provide answers to the questions that motivated this study:

(i) Are neural oscillations in place at birth? Yes. We found oscillations in the delta band to be active during resting state, and activity in all the frequency bands of interest (delta, theta, alpha, beta, low-gamma bands, high-gamma band) to be modulated during speech stimulation.

(ii) Do neural oscillations relate to speech processing at birth? Yes. Based on the comparisons between neural oscillations during speech stimulation and resting state, as well as the pairwise comparisons of the language conditions, we claim that neural oscillations in all the frequency bands of interest (delta, theta, alpha, beta, low-gamma bands, high-gamma band) are modulated (even if only weakly) during speech processing at birth. More specifically, we showed that theta and low-gamma oscillations seem to be involved in the recognition of the native language, and the discrimination of rhythmically different languages.

(iii) Does familiarity with the prenatally heard language have an effect on the behaviour of neural oscillations? Yes. We found that neural oscillations had higher overall activation for the familiar language (French) than for the rhythmically different unfamiliar one (English).

(iv) Do neural oscillations relate to the ability newborns have to discriminate rhythmically different languages? Yes. Activity in the theta and low-gamma bands seem to reflect rhythmic discrimination, exhibiting higher activation for the syllable-timed languages (French and Spanish) that for the stress-timed language (English).

(v) Is there synchronization between neural oscillations and the speech envelope (i.e. speech envelope tracking) at birth? Yes. We found that newborns are able to track the amplitude and phase of the speech envelope in familiar and unfamiliar languages equally well. We thus claimed that envelope tracking represents a basic auditory mechanism that takes place in the absence of comprehension and attention.

(vi) Is envelope tracking at birth shaped by prenatal experience? No. We found that envelope tracking results were not significantly different between the familiar and the unfamiliar languages, and thus claim that this ability is not shaped by prenatal experience.

CHAPTER 3: Neural oscillations in six-month-old infants

3.1 INTRODUCTION

At birth, infants can discriminate their native language (heard prenatally) from a rhythmically different one, but not from a rhythmically similar one (Nazzi et al., 1998; Ramus et al., 2000). It is only until a few months later that this within-rhythmic class discrimination develops (between 3.5 and 4 months of age, Bosch & Sebastián-Gallés, 1997; Molnar et al., 2013). One possible explanation for the acquisition of this within-rhythmic class discrimination ability is that infants might start to rely on the phoneme repertoire of their native language to be able to identify it (Molnar et al., 2013). Electrophysiological studies have shown that processing of the native language is characterized by an increase in the gamma band power, which is not present when processing a rhythmically similar unfamiliar language (6-month-olds, Peña et al., 2010; adults, Peña & Melloni, 2012). This in turn is consistent with the previous hypothesis, as the processing of (sub)phonemic units has been argued to be linked to the activity of gamma oscillations (Giraud & Poeppel, 2012).

Neural oscillations in other frequency bands have also been argued to be related to speech processing (theta oscillations to the processing of syllabic units, and delta oscillations to the processing of phrasal units, Giraud & Poeppel, 2012). However, developmental studies exploring neural oscillations during speech processing are scarce, and little is known about the changes that oscillations have as language acquisition develops in the infant brain. To shed some light on this growing field, we investigated neural oscillations and their relationship with speech processing in 6-month-old French learning infants with the aim to answer the following questions:

i. How do neural oscillations relate to between-rhythmic class discrimination (French vs English, Spanish vs English) at 6 months?
ii. How do neural oscillations relate to within-rhythmic class discrimination (French vs Spanish) at 6 months?
iii. How do neural oscillations change after the first 6 months of life?
iv. How do language experience and age affect the synchronization of neural oscillations to the speech envelope (speech envelope tracking)?

To tackle these developmental questions, we tested 6-month-old French learning infants in three conditions while recording their neural activity with EEG: (1) while listening to utterances in their native language (French), (2) a rhythmically similar unfamiliar language (Spanish), (3) and in a rhythmically different unfamiliar language (English). By assessing

how neural oscillations behave in the previous conditions, and comparing their behaviour to the results at birth (Chapter 2), we aim to explore how neural oscillations change in the infant brain after language experience and additional brain maturation take place.

It needs to be noted that even though the within-rhythmic class discrimination ability is present around 4 months of age, here we looked at 6-month-olds because we considered it would be beneficial for data quality to test infants when they had already developed head control, and were able to hold their heads up without too much movement during the EEG recording session.

The three tested conditions were part of one single experiment, and each participant heard the three languages (English, French and Spanish). In order to address the various questions of our study, we analysed the collected data applying different methodologies for each question: time-frequency analysis for neural oscillations; and cross-correlation and phase-coherence for speech envelope tracking (as done in the newborn study).

3.2 MATERIALS AND METHODS

3.2.1 Participants

For this experiment, infants were recruited in two different ways: one portion of the group corresponds to participants who took part of the newborn experiment and were re-contacted (n=17), while the other portion corresponds to participants who were contacted through the database of our Babylab, and who were tested for the very first time (n=19). All parents gave written informed consent prior to participation, and were present during the testing session. The protocol for this study was approved by the CER Paris Descartes ethics committee of the Paris Descartes University. The inclusion criteria were: i) being full-term and healthy, ii) being 6 months old ± 15 days, and iii) being exposed to French at least 80% of the time since birth (according to parental report). Some exceptions regarding the age at the time of the experiment were allowed when testing the participants who were re-contacted from the newborn group. Due to availability constraints from the families, and our strong interest in collecting this longitudinal data, 4 participants were tested outside of the desired age range: one participant at 5 months, one participant at 10 months, and two participants at 11 months. A total of 36 full-term, healthy infants were tested, however, 11 participants were excluded from data analysis due to: bad data quality resulting in an insufficient number of non-rejected trials (n = 4), not finishing the experiment due to crying (n = 5), or technical problems (n = 2).

Thus, electrophysiological data from 25 infants (age 204 days ± 48 d; range 152-340 d; 10 girls, 15 boys) were included in the analysis (12 longitudinal, 13 cross-sectional).

3.2.2 Procedure

The procedure was very similar to the one implemented in the experiment with newborns (Chapter 2). We tested infants by presenting them with naturally spoken sentences in three languages while simultaneously recording their neural activity using electroencephalography (EEG). In order to avoid motion artifacts, we also presented a silent video of a children's animated cartoon, that was reproduced on a monitor placed in front of the participants during the entire recording session (as it was done in Abrams et al., 2008). The EEG recordings were conducted at two different locations: one was a dimmed, quiet room at the Robert-Debré Hospital in Paris, were participants were tested at birth; and the other one was a sound-attenuated testing booth at the Babylab of the Université Paris Descartes. The reason for having two testing locations was to accommodate to the constraints of the re-contacted families (longitudinal) who were unable to visit the Babylab (n=10).

Participants heard one of three possible sets of sentences: 10 infants heard set1, 14 infants heard set2, and 1 infant heard set3. Considering that some of the participants were tested at birth, we controlled for the set of sentences that was presented to them at 6 months, to ensure that they did not hear the same stimuli and rule out that any memory trace would have an effect on the new recordings. During the recording session, infants were seated on a parent's lap, and approximately 60 cm away from the monitor that was placed in front of them. The stimuli were delivered bilaterally through two loudspeakers positioned on each side of the monitor using the experimental software E-Prime. The sound volume was set to a comfortable conversational level (~65-70 dB).

Following the procedure implemented in the newborn experiment (Chapter 2), we presented participants with one sentence per language, and repeated it 50 times to ensure sufficiently good data quality. For this experiment we used fewer repetitions than for the newborn group because 6-month-olds were awake, and we anticipated that they were not going to sit still for a sufficiently long period of time. We also excluded the resting state condition from this experiment due to the aforementioned time and attentional constraints. The experiment consisted of 3 blocks, each block containing the 50 repetitions of the test sentence in a given language, each block thus lasted around 3.5 minutes. An interstimulus interval of random duration (between 1 – 1.5 s) was introduced between sentence repetitions,

and an interblock interval of 10 s was introduced between language blocks. The entire recording session lasted about 11 minutes.

To ensure that the within-rhythmic class discrimination could be investigated, we presented French and Spanish in the first two blocks, to guarantee that we could collect data for these two conditions even when infants would fail to finish the experiment (e.g. due to being tired or crying towards the end). However, the order of the first two blocks (French and Spanish) was pseudo-randomized and approximately counter-balanced across participants. The English block was always presented at the end of the experiment.

3.2.3 Stimuli

We tested 6-month-olds in the following three languages: their native language (French), a rhythmically similar unfamiliar language (Spanish), and a rhythmically different unfamiliar language (English). The stimuli that we used for the current experiment consisted of the same sentences used for the newborn experiment. See Table 2 (in Chapter 2) for sentence details.

3.2.4 EEG Data acquisition

EEG data were recorded with active electrodes and an acquisition system from Brain Products (actiCAP & actiCHamp). A 10-channel layout was used to acquire cortical responses from the following scalp positions: F7, F3, FZ, F4, F8, T7, C3, CZ, C4, T8 (see Fig. 6b in Chapter 2). The 10 channels correspond to the same topography that we implemented in the newborn experiment. Additional electrodes were placed on both mastoids for online reference, and a ground electrode was placed on the forehead. Data were recorded at a sampling rate of 500 Hz.

3.2.5 Envelope Tracking Analysis

To investigate whether neural oscillations of 6-month-olds synchronize to the envelope of a speech signal during a passive listening task, we compared the amplitude and the phase of these two oscillatory signals, in the same way as we did it for the newborn experiment (Chapter 2). We computed both measures to explore which features of the speech envelope are encoded by the infant brain. We pre-processed and analysed all the data using custom Matlab® scripts developed as part of this study.

Amplitude-Encoding Analysis

To assess amplitude-encoding of the speech envelope, we followed the same procedure described for newborns (Chapter 2): (i) band-pass filtered the EEG (1-40 Hz), (ii) epoched the trials, (iii) performed baseline correction, (iv) trial rejection, and (v) averaged non-rejected trials to obtain the evoked response. Each participant contributed on average 27 epochs for the English condition, 31 for the French and 31 for the Spanish. Participants who had less than 10 remaining epochs after epoch rejection were not included in the data analysis (n=4, as also reported in the Participants section).

The amplitude envelopes of the speech signals (low-pass filtered at 40 Hz) were compared to the EEG evoked responses by performing cross-correlation analysis (as detailed in Chapter 2). We then averaged the cross-correlation functions across the channels that were at least weakly correlated ($\rho>0.15$) in the three languages to obtain a mean time lag for each participant. As a last step, we quantified amplitude-encoding as the correlation between the evoked responses and the speech envelopes evaluated at the mean time lag obtained for each participant. The correlation values were Fisher transformed before statistical analysis, as is customary when performing statistical comparisons of correlation coefficients. Even though only weakly correlated channels were included when obtaining the participants' time lags, all channels were included in the amplitude-encoding correlation analysis.

Statistical Analyses of Amplitude-Encoding Results

To validate our correlation results and establish envelope tracking, first we assessed if their distributions were significantly different from zero by submitting them to one-sample t-tests (two-tailed). This was done for each language and each channel. As a second step, we also compared our amplitude-encoding results against chance, by implementing a permutation test that shuffled the EEG data 1000 times along the time dimension (as detailed in Chapter 2). The results obtained from the permuted data served as a baseline of comparison for our amplitude-encoding results. The p-values for the permutation test were derived as the number of permutations out of the 1000 whose correlation distributions were not significantly different from the correlation distributions obtained for the non-permuted data (in paired samples t-tests). As a third statistical analysis we tested linear mixed effects models to assess if envelope tracking was different across languages and/or across channels. Table 3 (in Chapter 2) presents the models that were run and compared in SPSS. The best fitting model

was chosen based on the AIC (Akaike information criterion) value. The sample size for all the analyses previously described was 25 participants.

Phase-Encoding Analysis

To assess phase-encoding of the speech envelope, we followed the same procedure described for newborns (Chapter 2): (i) band-pass filtered the EEG at the syllabic rate of the tested sentences (3-6 Hz), (ii) epoched the trials, (iii) performed baseline correction, (iv) trial rejection, and (v) averaged non-rejected trials to obtain the evoked response. The filtered EEG evoked responses (3-6 Hz) were compared to the filtered speech envelope (3-6 Hz) of the corresponding tested sentence by computing the phase-coherence between them (as detailed in Chapter 2).

Statistical Analyses of Phase-Encoding Results

To validate our phase-coherence results and establish phase-locking to the speech envelope, first we assessed if their distributions were significantly different from zero by submitting them to one-sample t-tests (two-tailed). This was done for each language and each channel. As a second step, we also compared our phase-encoding results against chance, by implementing a permutation test similar to the one used in the newborn experiment (Chapter 2). The phase-coherence results obtained from the permuted data served as a baseline of comparison against chance for the phase-encoding results. As a third statistical analysis we tested linear mixed effects models to assess if phase-locking was different across languages and/or across channels. Table 3 (in Chapter 2) presents the models that were run and compared in SPSS. The best fitting model was chosen based on the AIC (Akaike information criterion) value. The sample size for all the analyses previously described was 25 participants.

Developmental Differences in Envelope Tracking Results

To investigate whether language experience and developmental changes during the first 6 months of life have an effect on the ability to track the amplitude and the phase of the speech envelope, we compared the envelope tracking results at 6 months with those from the newborn group. By assessing how envelope tracking at 6 months differs from envelope tracking at birth when listening to familiar (French) and unfamiliar languages (Spanish and English), we were able to explore developmental changes that are general to speech

processing (present in the three languages) as well as changes specific to the acquired experience with the native language (present only for French).

We assessed the developmental differences of the amplitude-encoding and phase-encoding results by submitting the corresponding measures from both age groups to unpaired samples t-tests (two-tailed): i.e. correlations(0m) vs. correlations(6m); phase-coherence(0m) vs. phase-coherence(6m). This analysis was performed for each language and channel separately. The statistical comparison included 47 participants from the newborn group, and 25 participants from the 6-month-old group. We acknowledge that the different sample sizes from the two groups are a problem. To account for this, we performed two additional comparisons: one comparing 25 random participants from the newborn group to the 25 participants from the 6-month-old group (with a unpaired samples t-tests, two-tailed); and another one comparing 12 longitudinal participants (with a paired samples t-tests, two-tailed).

3.2.6 Time-Frequency Analysis

To investigate how neural oscillations behave during speech processing in the infant brain of 6-month-olds, we performed a time-frequency analysis on the EEG data. To do so, we followed the same procedure described for the newborn data (Chapter 2): (i) we removed the power line noise at 50 Hz, (ii) band-pass filtered the denoised EEG (1-100 Hz), (iii) epoched the trials, and (iv) performed trial rejection. Each participant contributed on average 27 epochs for the English condition, 31 for the French condition, and 31 for the Spanish condition. Participants who had less than 10 remaining trials after epoch rejection were not included in the data analysis. From the 25 participants who entered data analysis when assessing envelope tracking, 2 infants were rejected from the current time-frequency analysis because they did not provide a sufficient number of non-rejected trials (due to noise present in the 1-100 Hz range of the current analysis, but not present in the 1-40 Hz of the envelope tracking analysis). Therefore, a total of 23 participants were included in the time-frequency analysis.

We submitted the non-rejected epochs to a continuous wavelet transform using a Matlab® toolbox named ‘WTools’ (Parise & Csibra, 2013) to generate the time-frequency representations for each condition. Following the same steps described in the newborn experiment (Chapter 2), we obtained a response matrix of spectral amplitude values with 100 frequency bins (1-100 Hz) and 982 time samples (1,964 ms) for each language condition, per

participant. We then submitted the time-frequency response matrices to the following statistical analyses:

Analysis 1: Effect of Frequency & Language (6 frequency bands)

For this analysis we focused on the neural oscillations operating at 6 frequency bands of interest: delta, theta, alpha, beta, low-gamma, and high-gamma. Here we investigated how neural oscillations behaved across these 6 frequency bands and across languages. To carry out this analysis we followed the procedure described for newborns (Chapter 2) to obtain a mean spectral amplitude value for each frequency band of interest (delta, theta, alpha, beta, low-gamma, high-gamma). This analysis is thus similar to evaluating the spectrum of the EEG signal, specific to the 6 frequency bands of interest. To assess differences across frequency bands and languages, we submitted the mean spectral amplitude values from all the participants to a two-way repeated measures ANOVA with *Frequency-Band* and *Language* as the within-subjects factors.

Analysis 2: Effect of Language

Here we investigated whether neural oscillations operating in a broader frequency range (1-100 Hz) behave differently across the tested languages over the time course of stimulation. By assessing language differences, we aimed to explore which frequency bands (if any) reflect the familiarity that infants have with their native language (French), as well as the language discrimination that takes place between rhythmically similar (French, Spanish) and rhythmically different languages (French, Spanish, vs. English). We explored language differences in the neural oscillations by submitting the spectral amplitude values of the 100 frequency bins (1-100 Hz) over the 982 time samples from the 3 languages to a repeated measures ANOVA with language as the within-subjects factor. As post hoc analyses, we computed paired samples t-tests (two-tailed) to compare the languages pairwise.

Analysis 3: Developmental Differences (Newborns vs. 6-month-olds)

The objective of this analysis was to investigate the developmental changes that neural oscillations have after the first 6 months of life. By assessing how neural oscillations at 6 months of age differ from neural oscillations at birth when processing familiar (French) and unfamiliar languages (Spanish and English), we were able to explore two developmental

changes of different nature: (1) changes due to brain maturation and/or general cognitive development, which would be present in familiar and unfamiliar languages alike; (2) and changes due to the acquired experience with the native language, which would be present only for French (and possibly the rhythmically similar language, Spanish).

We assessed the developmental changes of neural oscillations by submitting the spectral amplitude values of the 100 frequency bins (1-100 Hz) over the 982 time samples of the two age groups (newborns and 6-month-olds) to unpaired samples t-tests (two-tailed). This analysis was performed for each language separately. This statistical comparison included 47 participants from the newborn group, and 23 participants from the 6-month-old group. We acknowledge that the different sample sizes from the two groups are a problem. To account for this, we performed two additional comparisons: one comparing 23 random participants from the newborn group to the 23 participants from the 6-month-old group (with a unpaired samples t-tests, two-tailed); and another one comparing 10 longitudinal participants (with a paired samples t-tests, two-tailed).

3.3 RESULTS

This section presents the results of the various analyses performed on the EEG data from the 6-month-old infants.

3.3.1 Envelope Tracking

To assess speech envelope tracking, we compared the EEG signals to the speech envelopes using two measures: amplitude-encoding and phase-encoding. For illustration, the average EEG time series from one participant are displayed in Figure 20.

Amplitude-Encoding Results

Figure 21 presents the envelope tracking results of a single participant as illustration. In this figure we can observe how the EEG recordings (red, blue and green curves) track the speech envelopes (black curves) in the three languages.

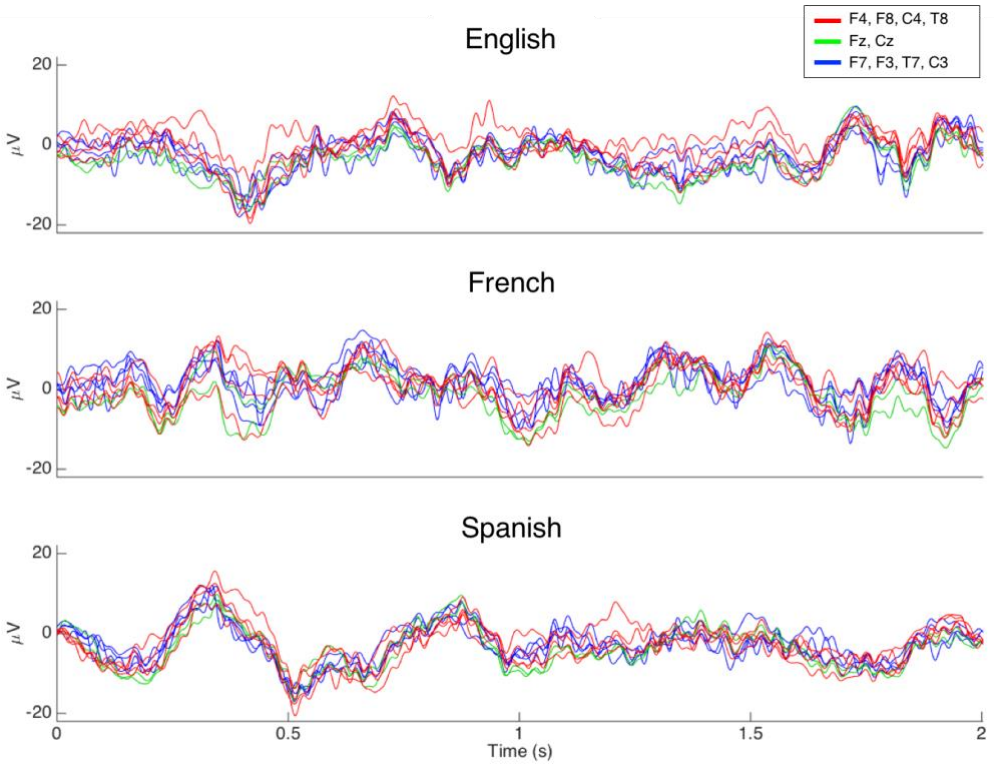


Figure 20. Average EEG time series from one 6-month-old participant for the three language conditions. The red curves correspond to channels in the right hemisphere (F4, F8, C4, T8); the green curves to channels in the central region (Fz, Cz); and the blue curves to channels in the left hemisphere (F7, F3, T7, C3). These EEG responses are the average of the non-rejected trials.

Mean correlation values across the whole group of infants varied from channel to channel, ranging from 0.094 to 0.217 (mean 0.168; SD 0.0339) for **English**, from 0.082 to 0.149 (mean 0.117; SD 0.0201) for **French**, and from 0.068 to 0.130 (mean 0.092; SD 0.0199) for **Spanish**. These group results are listed in Table 7 and displayed in Fig. 22 (yellow boxes) per channel and per language. The correlation distributions from all channels in the English and French condition were significantly higher than zero when submitted to one-sample t-tests (two-tailed). However, only channels F3 and Fz were significantly higher than zero for the Spanish condition. The p-values from these t-tests were corrected for multiple comparisons (i.e. to adjust the family-wise error rate for each comparison) using the Bonferroni correction, and their details are listed in Table 7.

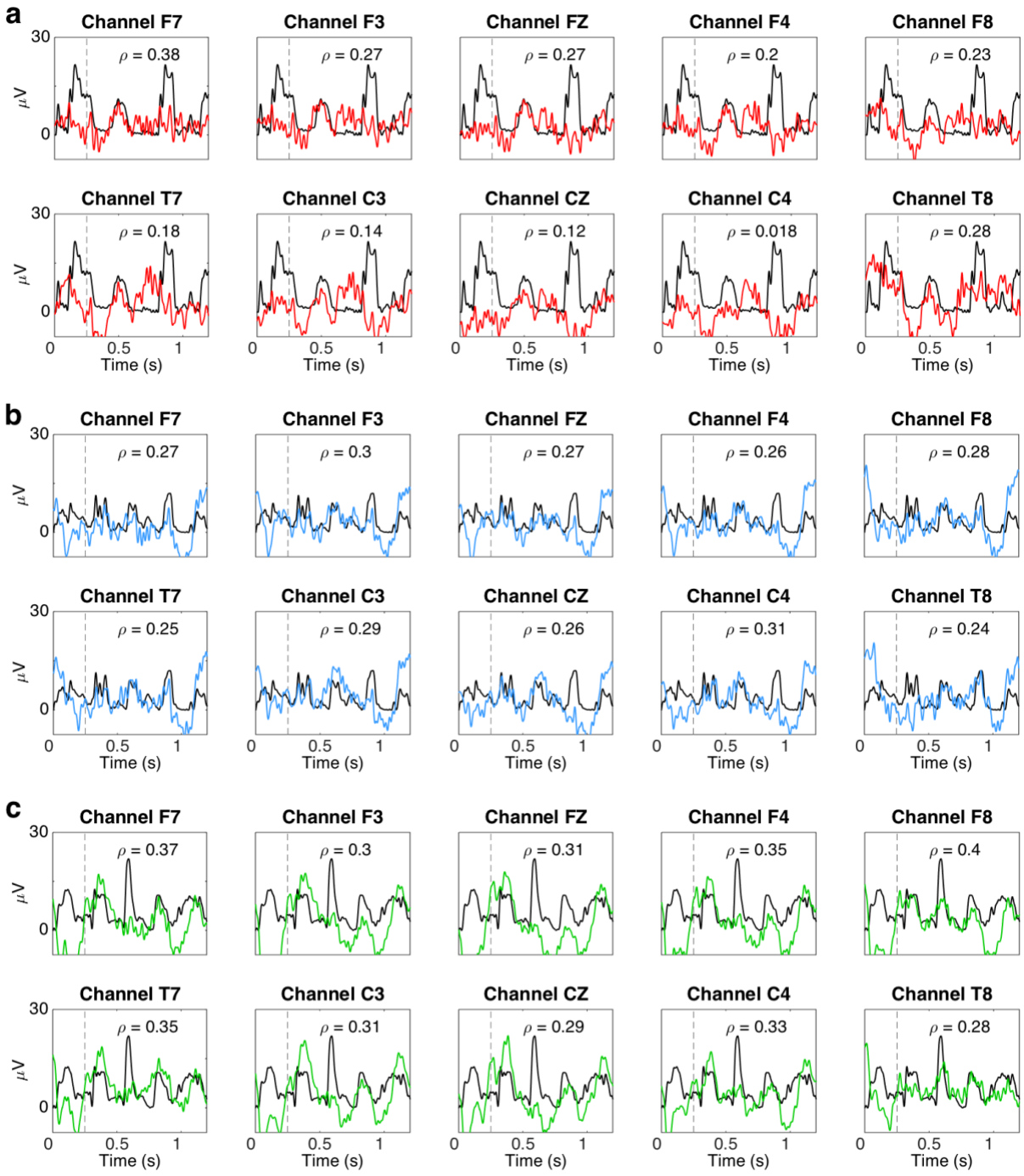


Figure 21. Amplitude envelope tracking of speech in an individual 6-month-old participant. **a)** Envelope-tracking of an English sentence. The black curves represent the speech envelope and the red curves the cortical activity measured per channel. For visualization purposes the EEG data was shifted backward in time; this time shift corresponds to the time lag between the neural response and the speech signal. The Spearman correlation between the two signals is indicated by "ρ". **b)** Envelope-tracking of a French sentence. The blue curves represent the cortical activity measured per channel. **c)** Envelope-tracking of a Spanish sentence. The green curves represent the cortical activity measured per channel. The vertical line at 250 ms indicates the beginning of the envelope-following period.

	ENGLISH			FRENCH			SPANISH		
Channel	mean	p-value	t-value	mean	p-value	t-value	mean	p-value	t-value
F7	0.164	0.00	6.007	0.117	0.02	3.402	0.105	0.06	3.036
F3	0.165	0.00	5.536	0.139	0.00	4.427	0.130	0.01	3.786
Fz	0.186	0.00	6.519	0.120	0.01	3.744	0.106	0.02	3.505
F4	0.189	0.00	6.752	0.133	0.00	5.286	0.097	0.12	2.728
F8	0.152	0.00	7.627	0.108	0.02	3.378	0.103	0.12	2.704
T7	0.140	0.00	4.442	0.082	0.03	3.264	0.068	0.69	1.902
C3	0.186	0.00	5.421	0.119	0.00	4.662	0.089	0.14	2.655
Cz	0.217	0.00	8.198	0.098	0.03	3.299	0.075	0.44	2.130
C4	0.186	0.00	7.358	0.149	0.00	5.917	0.069	0.77	1.847
T8	0.094	0.03	3.355	0.103	0.01	3.717	0.076	0.33	2.266

Table 7. Group mean correlation values for the 6-month-old participants per channel and per language. The correlation distributions were submitted to one-sample t-tests (two-tailed) to compare them to zero. The p-values (Bonferroni corrected) and t-values of these t-tests are presented here per channel. Non-significant results are highlighted in red.

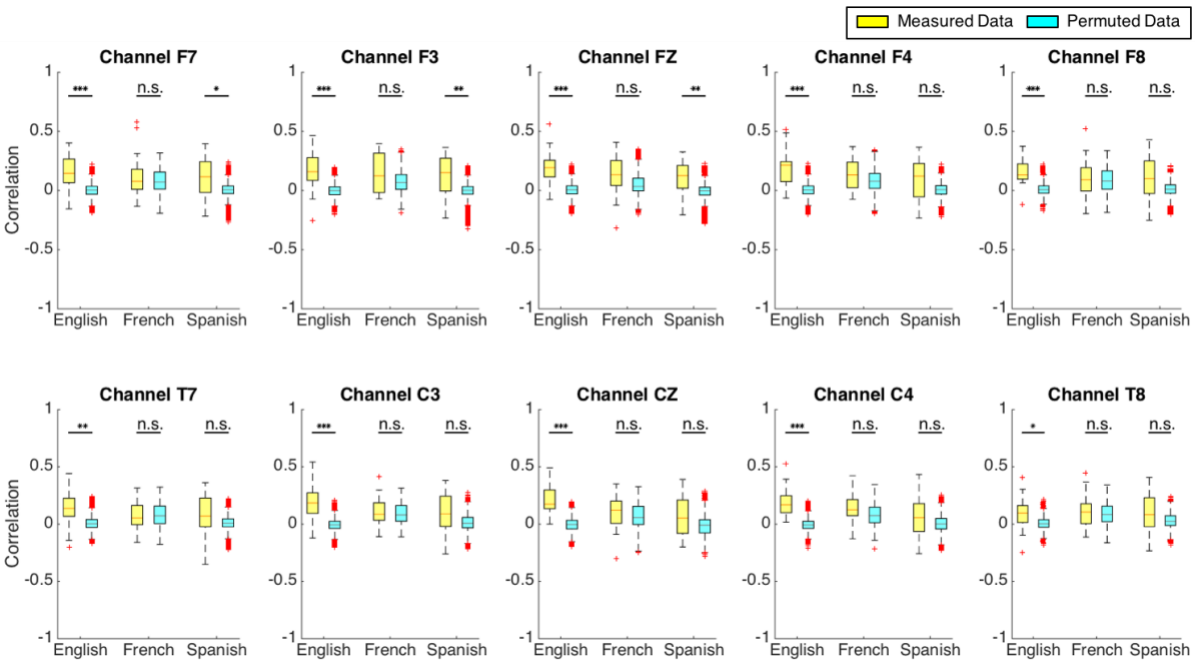


Figure 22. Amplitude envelope tracking results for 6-month-old participants are displayed for each channel, and for the three languages in a box and whisker plot. The yellow boxes represent the distributions of the results from the measured EEG data, and the blue boxes represent the distributions of the results from the permuted data. n.s.: non-significant, * p<0.05, ** p<0.01, *** p<0.001.

To further establish envelope tracking, we used a permutation test to compare against chance. We permuted the EEG data 1000 times, and performed the envelope tracking analysis at each permutation. The results from the permutation test are displayed in Fig. 22 (blue boxes) next to the results from the measured data (yellow boxes). As it can be seen in Fig. 22, only the measured correlation distributions from **English** were significantly higher than chance for all channels. For the **French** condition none of the correlation distributions was higher than chance, while for **Spanish** only channels F7, F3, Fz exhibited significant results. These results show that the brain of 6-month-olds encodes the amplitude of the speech envelope for the rhythmically different unfamiliar language (English) but not that of the native language (French). Additionally, amplitude-encoding of the rhythmically similar unfamiliar language (Spanish) seems to be weak.

To assess possible differences in the amplitude-encoding results across languages and channels, we implemented linear mixed effects models with the correlation values as the dependent variable (see Methods). The best fitting model (model '*m2'*, Table 3 in Chapter 2) had *Language* as fixed factor, and *Participant* and *Language* as random factors. For this model, the main effect of the fixed factor *Language* was marginally significant [$F(2,72) = 2.841$, $p = 0.065$].

Developmental Differences in Amplitude-Encoding Results

In this analysis we assessed how envelope tracking by means of amplitude-encoding changed after the first 6 months of life. Figure 23 displays the amplitude-encoding differences between the 6-month-old data and the newborn data when including all the participants for both groups (*comparison_1*): 47 participants from the newborn group, and 25 participants from the 6-month-old group. Additionally, we also performed two balanced comparisons including subsets of the data. Figure 24 displays the amplitude-encoding age differences when including 25 random participants from the newborn group, and the 25 participants from the 6-month-old group (*comparison_2*). Figure 25 displays the comparison results when including 12 longitudinal participants (*comparison_3*). These comparisons give some indication of English amplitude-encoding results being higher at 6 months than at birth, with (3/10) channels showing this trend in *comparison_1*, (6/10) channels in *comparison_2*, and (5/10) channels in *comparison_3*. For French, no significant differences were found between the correlation results of both age groups, and for Spanish, only (1/10) channel had significantly lower correlation results at 6 months compared to birth (*comparison_2* and *comparison_3*).

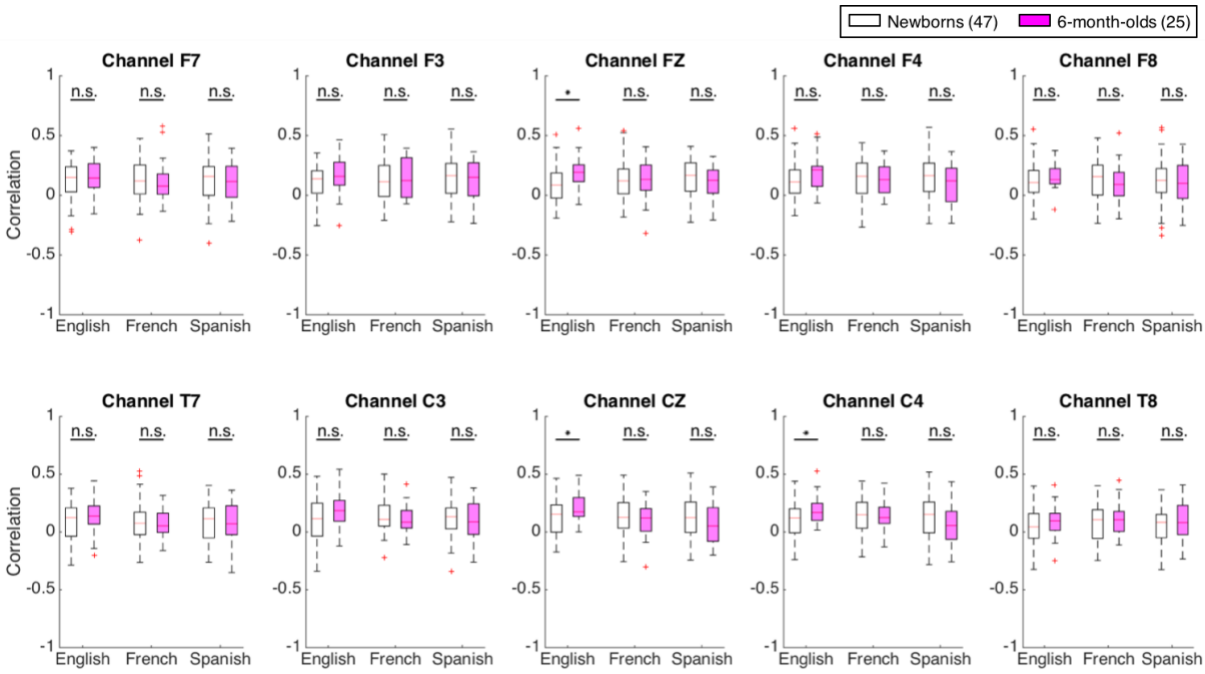


Figure 23. Correlation distributions at birth (white) compared to correlation distributions at 6 months (magenta) when including all the participants for both groups (47 newborns and 25 six-month-olds). Correlation results are displayed for each channel, and for the three languages in a box and whisker plot. n.s.: non-significant, * $p<0.05$.

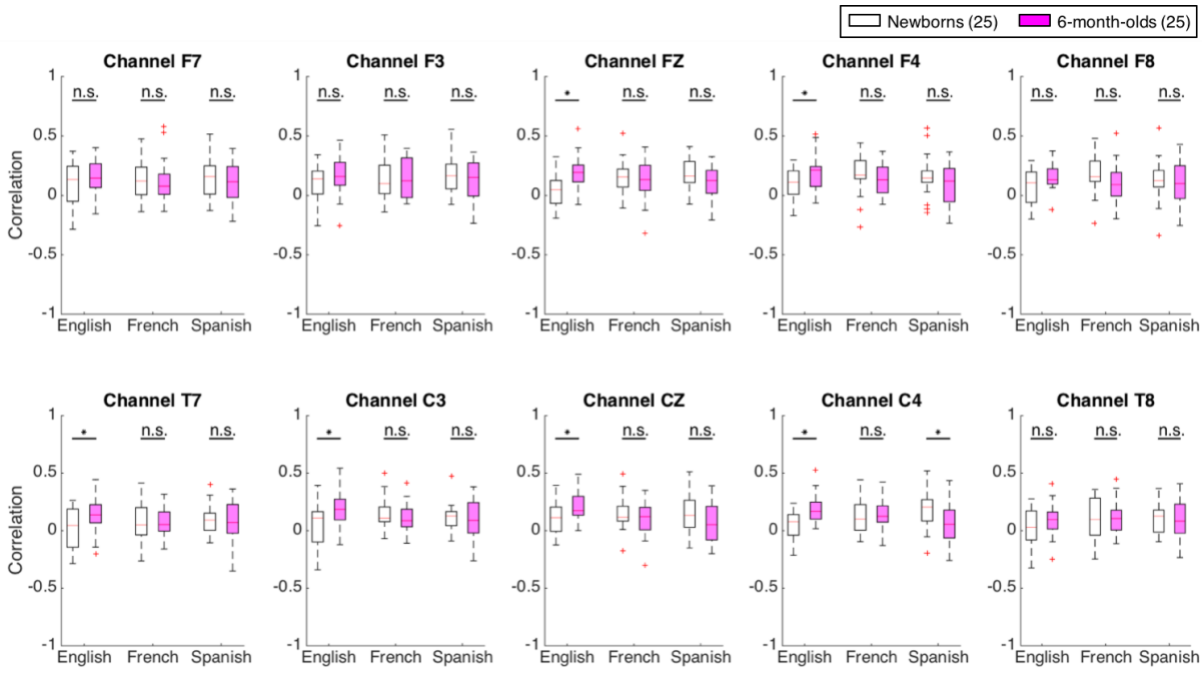


Figure 24. Correlation distributions at birth (white) compared to correlation distributions at 6 months (magenta) when including 25 random newborn participants and 25 participants from the 6-month-old group. Correlation results are displayed for each channel, and for the three languages in a box and whisker plot. n.s.: non-significant, * $p<0.05$.

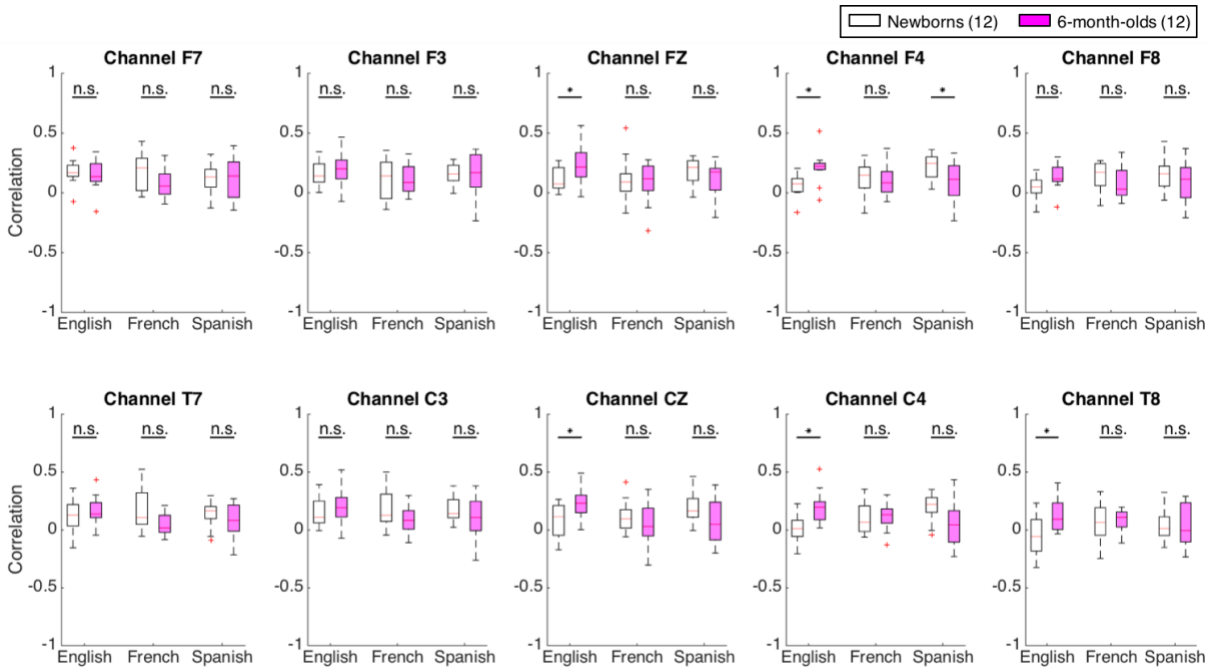


Figure 25. Correlation distributions at birth (white) compared to correlation distributions at 6 months (magenta) when including only the longitudinal participants (12). Correlation results are displayed for each channel, and for the three languages in a box and whisker plot. n.s.: non-significant, * $p<0.05$.

Phase-Encoding Results

Mean phase-coherence values across the whole group of infants varied from channel to channel, ranging from 0.276 to 0.336 (mean 0.312; SD 0.0181) for **English**, from 0.253 to 0.351 (mean 0.288; SD 0.0288) for **French**, and from 0.243 to 0.342 (mean 0.300; SD 0.0295) for **Spanish**. These group results are listed in Table 8 and displayed in Fig. 26 (yellow boxes) per channel and per language. The phase-coherence distributions were significantly higher than zero for all channels and languages, when submitted to one-sample t-tests (two-tailed). The p-values from these t-tests were corrected for multiple comparisons (i.e. to adjust the family-wise error rate for each comparison) using the Bonferroni correction, and their details are listed in Table 8.

To further establish phase-encoding we used a permutation test, similar to the one implemented for the amplitude-encoding results, to compare against chance. The results from the permutation test are displayed in Fig. 26 (blue boxes) next to the results from the measured data (yellow boxes). Our phase-coherence results are significantly higher than the permutation results for all the channels and all the languages. These results thus establish that

the brain of 6-month-olds encodes the phase of the speech envelope in all three languages more accurately than what would be expected purely by chance.

	ENGLISH			FRENCH			SPANISH		
Channel	Mean	p-value	t-value	mean	p-value	t-value	mean	p-value	t-value
F7	0.276	0.00	8.392	0.253	0.00	9.326	0.243	0.00	9.793
F3	0.308	0.00	8.559	0.280	0.00	9.822	0.273	0.00	8.893
Fz	0.312	0.00	10.689	0.274	0.00	10.015	0.288	0.00	9.251
F4	0.336	0.00	9.816	0.260	0.00	9.269	0.296	0.00	8.741
F8	0.323	0.00	9.678	0.268	0.00	11.338	0.322	0.00	10.343
T7	0.316	0.00	9.469	0.289	0.00	8.555	0.289	0.00	10.062
C3	0.301	0.00	8.951	0.309	0.00	12.556	0.342	0.00	9.918
Cz	0.313	0.00	10.162	0.307	0.00	9.927	0.333	0.00	12.340
C4	0.296	0.00	8.455	0.292	0.00	9.632	0.297	0.00	9.932
T8	0.335	0.00	8.763	0.351	0.00	11.399	0.313	0.00	8.046

Table 8. Group mean phase-coherence values for the 6-month-old participants per channel and per language. The phase-coherence distributions were submitted to one-sample t-tests (two-tailed) to compare them to zero. The p-values (Bonferroni corrected) and t-values of these t-tests are presented here per channel. All channels were significantly higher than zero for the three languages.

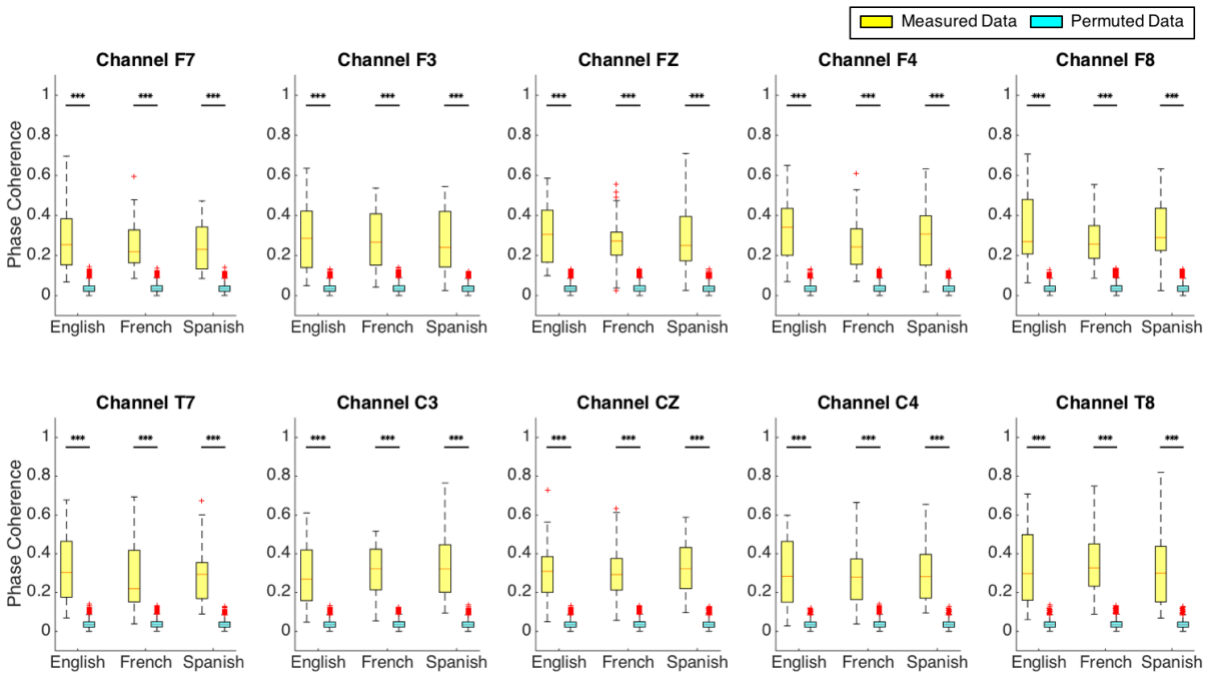


Figure 26. Phase-coherence results for 6-month-old participants are displayed for each channel, and for the three languages in a box and whisker plot. The yellow boxes represent the distributions of the results from the measured EEG data, and the blue boxes represent the distributions of the results from the permuted data. *** $p<0.001$.

To assess possible differences in the phase-encoding results across languages and channels, we implemented linear mixed effects models with the phase-coherence values as the dependent variable (see Methods). The best fitting model (model '*m2'*, Table 3 in Chapter 2) had *Language* as fixed factor, and *Participant* and *Language* as random factors. For this model, there was no main effect from the fixed factor *Language* [$F(2,74) = 0.392$, $p = 0.677$].

Developmental Differences in Phase-Encoding Results

In this analysis we assessed how envelope tracking by means of phase-encoding changed after the first 6 months of life. Figure 27 displays the phase-encoding differences between the 6-month-old data and the newborn data when including all the participants for both groups (*comparison_1*): 47 participants from the newborn group, and 25 participants from the 6-month-old group. Additionally, we also performed two balanced comparisons including subsets of the data. Figure 28 displays the phase-encoding age differences when including 25 random participants from the newborn group, and the 25 participants from the 6-month-old group (*comparison_2*). Figure 29 displays the comparison results when including 12 longitudinal participants (*comparison_3*). These comparisons show that phase-encoding results are lower at 6 months than at birth in only one or two channels for each language: For English (1/10) channel in *comparison_2*. For French, (2/10) channels in *comparison_1* and *comparison_2*, and (1/10) channel in *comparison_3*. For Spanish, (2/10) channels in *comparison_1*, and (1/10) channel in *comparison_2*.

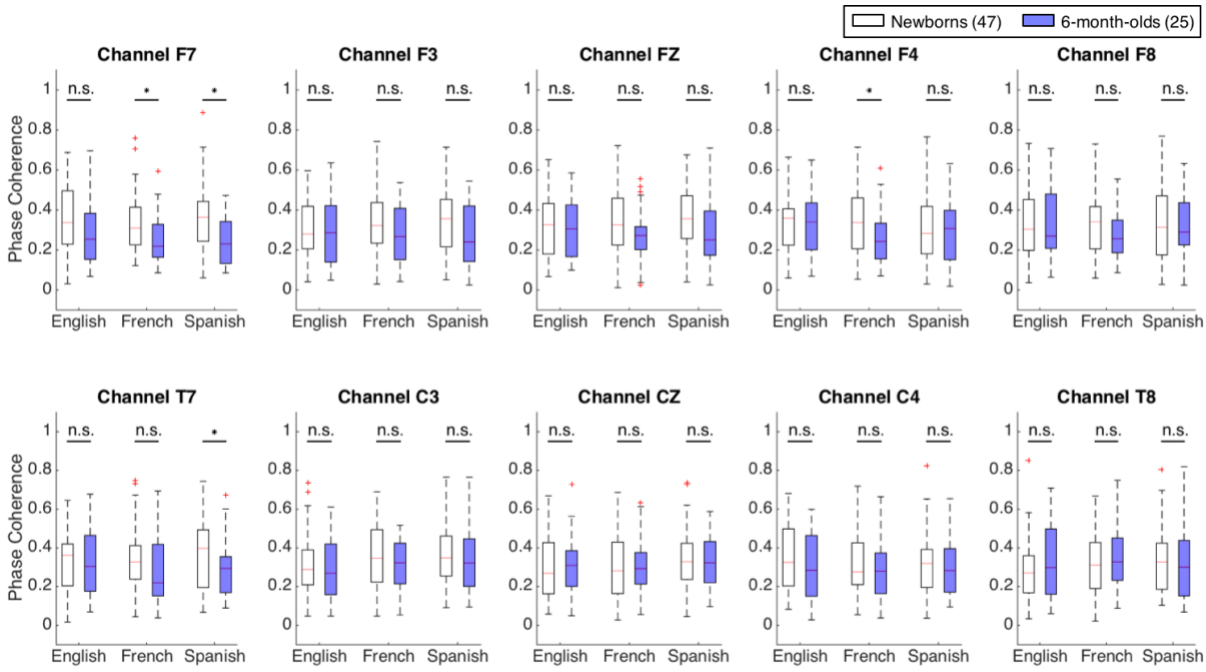


Figure 27. Distributions of phase-coherence results at birth (white) compared to at 6 months (purple) when including all the participants for both groups (47 newborns and 25 six-month-olds). Phase-coherence results are displayed for each channel, and for the three languages in a box and whisker plot. n.s.: non-significant, * $p<0.05$.

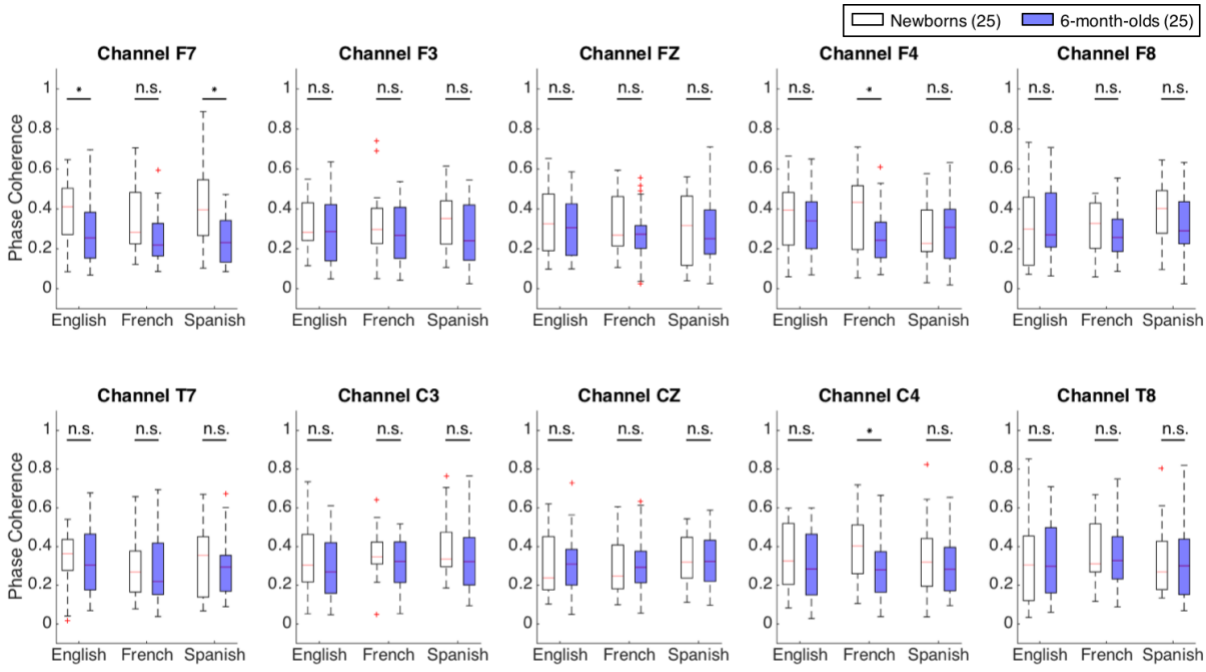


Figure 28. Distributions of phase-coherence results at birth (white) compared to at 6 months (purple) when including 25 random newborn participants and the 25 participants from the 6-month-old group. Phase-coherence results are displayed for each channel, and for the three languages in a box and whisker plot. n.s.: non-significant, * $p<0.05$.

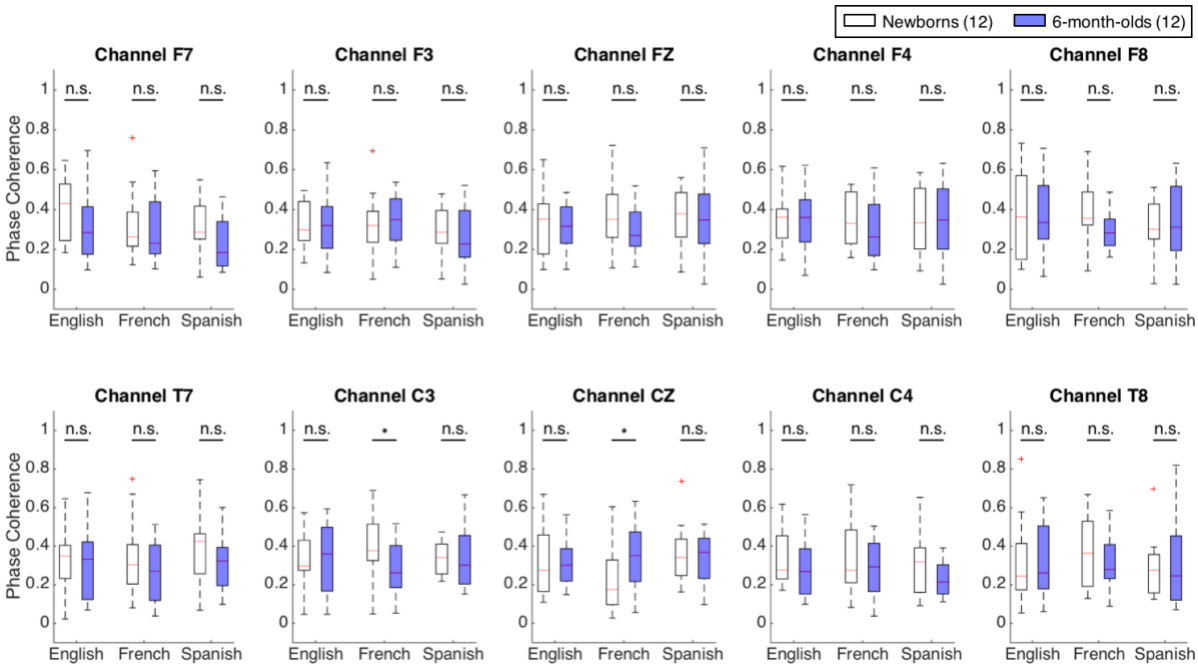


Figure 29. Distributions of phase-coherence results at birth (white) compared to at 6 months (purple) when including only the longitudinal participants (12). Correlation results are displayed for each channel, and for the three languages in a box and whisker plot. n.s.: non-significant, * $p<0.05$.

3.3.2 Time-Frequency Analysis

In order to investigate the behaviour of neural oscillations in the brain of 6-month-old infants, we obtained the time-frequency responses for the three tested languages (English, French, Spanish), and submitted them to various statistical analyses to explore differences across frequencies and languages, as well as along time. Group mean time-frequency responses are displayed in Figure 30. In this section we will be plotting the spectral amplitude of the time-frequency responses (the square root of the power).

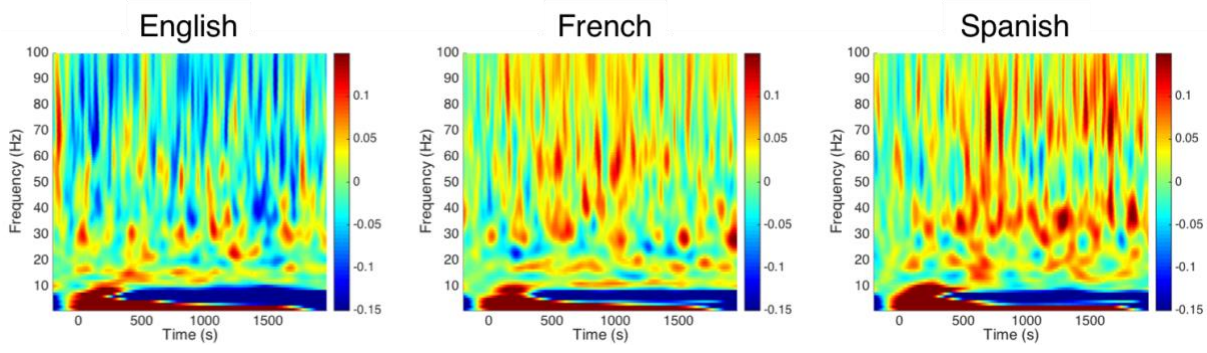


Figure 30. Group mean time-frequency responses for the three tested languages: English, French, and Spanish. Each plot is the average of the non-rejected epochs, the 10 channels and the 23 participants who were included in this analysis. The color bar on the right of each plot shows the spectral amplitude scale.

Analysis 1: Effect of Frequency & Language (6 frequency bands)

Figure 31 displays the mean spectral amplitude values per frequency band grouped per language condition. These plots are similar to the spectrum of the EEG signal, specific to the 6 frequency bands of interest. To investigate if the spectral amplitude values were different across languages and/or across frequency bands we applied a two-way repeated measures ANOVA with *Frequency-Band* and *Language* as the within-subjects factors. We found a main effect of *Frequency-Band* [$F(1.531,33.675) = 37.016$, $p < 0.0001$], and *Language* [$F(2,44) = 6.693$, $p = 0.003$], and a marginal interaction [$F(4.197,92.326) = 2.174$, $p = 0.075$]. From pairwise comparisons (Bonferroni corrected) we found that the *Frequency-Band* effect was due to the **delta band** being significantly higher than all the other bands, and the **theta band** being significantly lower than all the other bands. Similarly, the *Language* effect was due to **English** being significantly lower overall across all frequency bands than **French** and **Spanish**.

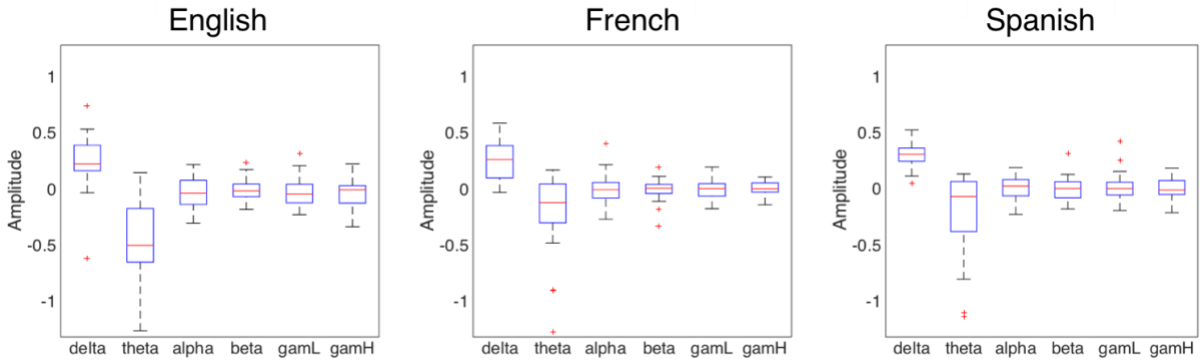


Figure 31. Mean spectral amplitude values for the six frequency bands of interest (delta, theta, alpha, beta, low-gamma, high-gamma) per language condition (English, French, Spanish). Activity in the delta band is significantly higher than in all the other frequency bands, and activity in the theta band is significantly lower than in the other bands. Activation for English was lower than for French and Spanish (Analysis 1).

Analysis 2: Effect of Language

In this analysis we assessed how neural oscillations differ across the three tested language conditions (English, French, Spanish) over the time course of stimulation. Figure 32a displays a time-frequency matrix with the results from the one-way repeated measures ANOVA that had language as the within-subjects factor (uncorrected for multiple comparisons). The red regions represent the time-frequency windows where the amplitude of the neural oscillations

was significantly different across the 3 languages. Language differences were found across the whole spectrum, with some bands displaying a stronger effect than others: the language effect is strongest in the delta, theta and high-gamma bands.

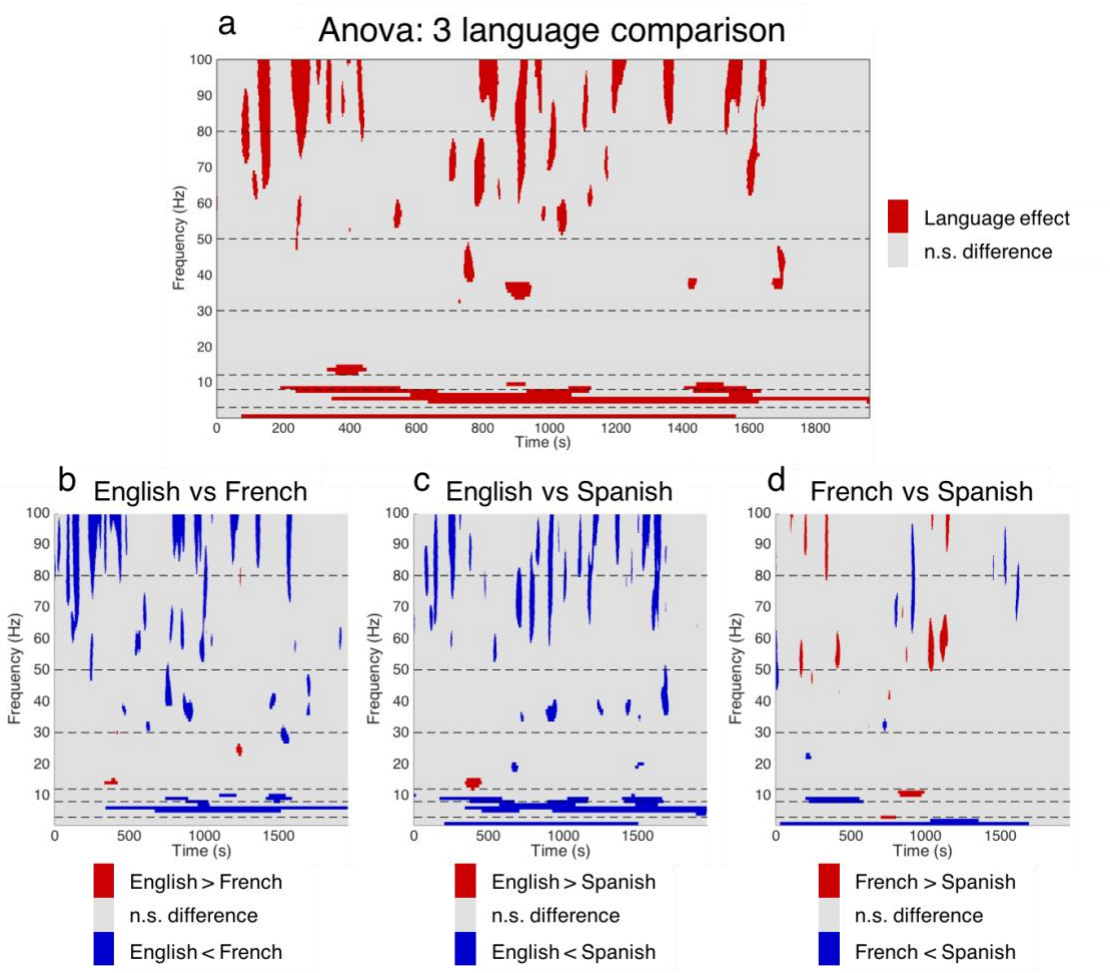


Figure 32. Time-frequency responses of significant language effect. **a)** Results of an ANOVA test between the neural responses of the 3 languages. The red regions represent the time-frequency windows where the amplitude of the neural oscillations was significantly different across the 3 languages. **b)** Results of a t-test between the neural responses of English and French. c) Results of a t-test between the neural responses of English and Spanish. D) Results of a t-test between the neural responses of French and Spanish. The coloured regions represent the time-frequency windows where the amplitude of the neural oscillations was significantly different across the 2 languages (Analysis 2).

To further explore the origin of these language differences we compared each pair of languages applying paired samples t-tests. Figure 32 (b,c,d) displays the time-frequency differences for each pairwise comparison. From these comparisons we can point out the following observations:

- Differences in the **delta** band (1-3 Hz) are only present in the comparisons that include Spanish.
- Differences in the **theta** band (4-8 Hz) are only present in the comparisons between rhythmically different languages (English vs. French, and English vs. Spanish).
- Small differences in the **alpha** band (9-12 Hz) are present in all the comparison (in restricted time windows).
- Only very small differences are present in the **beta** band (13-30 Hz).
- Differences in the **low-gamma** band (31-50 Hz) are mostly present in the comparisons between rhythmically different languages (English vs. French, and English vs. Spanish).
- Differences in the **high-gamma** band (50-100 Hz) are present in all the comparison. However, differences are more salient in the comparisons between rhythmically different languages (English vs. French, and English vs. Spanish). Importantly, when comparing rhythmically similar languages, the familiar language (French) showed greater activation than the unfamiliar language (Spanish).

Analysis 3: Developmental Differences (Newborns vs. 6-month-olds)

In this analysis we assessed how neural oscillations changed after the first 6 months of life. Figure 33 displays the time-frequency differences between the 6-month-old data and the newborn data when including all the participants for both groups: 47 participants from the newborn group, and 23 participants from the 6-month-old group. Differences were obtained using unbalanced unpaired samples t-tests (two-tailed). These results are uncorrected for multiple comparisons. From these comparisons we can point out the following observations:

- Developmental changes in the lower frequency bands (**delta, theta, alpha**) are similar for the three languages: Activation in these bands is higher for 6-month-olds during the first 500 ms (after stimulus onset), and lower afterwards.
- Developmental changes in the **beta** band are small, and present in the 3 languages (in restricted time windows). Activity for 6 month-olds is stronger than for newborns.
- Changes in the **gamma** band (low and high) are mainly present for the native language (French) and the rhythmically similar unfamiliar language (Spanish), with spectral amplitudes being higher for the 6 month-olds.

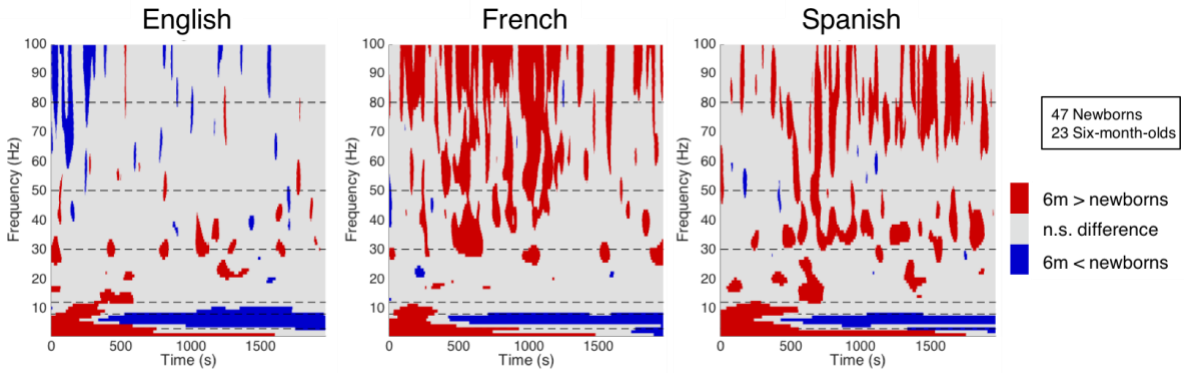


Figure 33. Time-frequency responses of significant developmental changes. Each matrix plot displays the time-frequency windows where the neural oscillations elicited by each language are significantly different between the two age groups (6-month-olds vs. newborns) when including all the participants for each groups (47 newborns and 23 six-month-olds). The red regions represent higher amplitude oscillations for the 6-month-olds, and the blue regions represent lower amplitude oscillations for 6-month-olds (Analysis 3).

Additionally, we also performed two balanced statistical comparisons including subsets of the data. Figure 34 displays the time-frequency differences between the 6-month-old data and the newborn data when including 23 random participants from the newborn group to match the number of participants from the 6-month-old group. Figure 35 displays the comparison results when including 10 longitudinal participants.

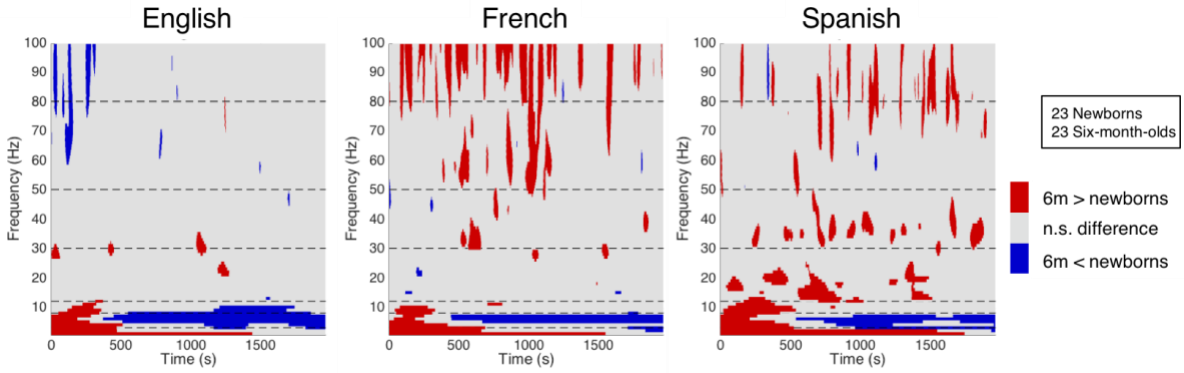


Figure 34. Time-frequency responses of significant developmental changes. Each matrix plot displays the time-frequency windows where the neural oscillations elicited by each language are significantly different between the two age groups (6-month-olds vs. newborns) when including 23 random participants from the newborn group to match the number of participants in the 6-month-old group. The red regions represent higher amplitude oscillations for the 6-month-olds, and the blue regions represent lower amplitude oscillations for the 6-month-olds (Analysis 3).

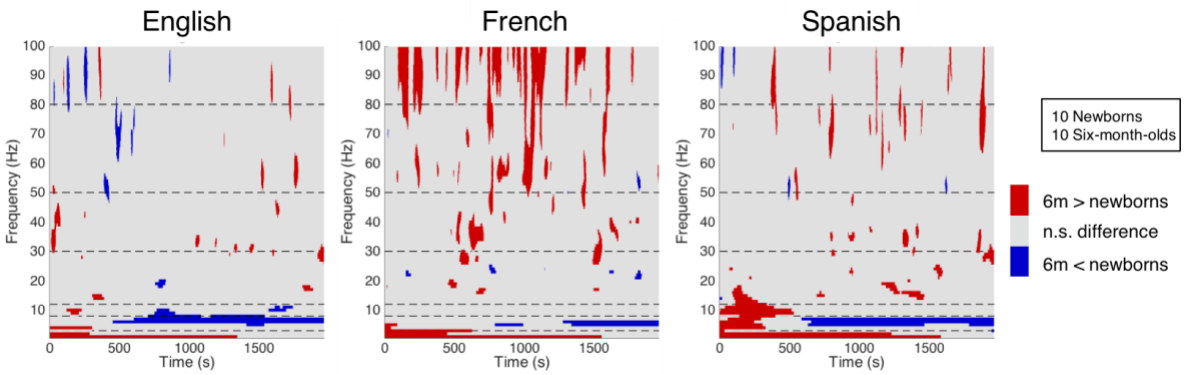


Figure 35. Time-frequency responses of significant developmental changes. Each matrix plot displays the time-frequency windows where the neural oscillations elicited by each language are significantly different between the two age groups (6-month-olds vs. newborns) when including 10 longitudinal participants. The red regions represent higher amplitude oscillations for the 6-month-olds, and the blue regions represent lower amplitude oscillations for the 6-month-olds (Analysis 3).

In these balanced comparisons we can observe the same age effect that was found for the unbalanced comparison: developmental changes are similar for the three languages in the lower frequency bands (delta, theta, alpha), while changes in the gamma band are mainly present for the native language (French) and the rhythmically similar one (Spanish).

Summary of Time-Frequency Results

Table 9 summarizes the results from all the time-frequency analyses presented in this chapter.

Analysis	Type of data	Effect Tested	Results
1	6 frequency bands	Frequency & Language	delta > other bands theta < other bands English < French, Spanish
2	Full-band time-frequency data	Language	rhythmic discrimination is strongest in the theta and gamma (30-100 Hz) bands: English < French, Spanish within-rhythmic class discrimination is seen in the high-gamma band (50-100 Hz): French > Spanish
3	Full-band time-frequency data	6-month-olds vs. Newborns	delta: 6m > 0m theta: 6m > 0m (beginning) 6m < 0m (after 500 ms) alpha: 6m > 0m (beginning) 6m < 0m (after 1000 ms, English) beta: 6m > 0m low-gamma: 6m > 0m high-gamma: 6m > 0m (French, Spanish) 6m < 0m (English)

Table 9. Summary of results for the time-frequency analyses.

3.4 DISCUSSION

In this chapter we investigated neural oscillations in 6-month-old infants in 3 language conditions: (1) while listening to sentences in their native language, (2) in a rhythmically similar unfamiliar language, and (3) in a rhythmically different unfamiliar language. By exploring how the neural oscillations behave in these different languages we aimed to investigate: (i) how neural oscillations relate to between-rhythmic class discrimination, (ii) and within-rhythmic class discrimination, (iii) how neural oscillations change after the first 6 months of life, (iv) and whether the synchronization of neural oscillations to the speech envelope is still present at this age. To tackle these questions, we analysed the infant EEG data using various methods: speech envelope tracking by means of amplitude-encoding and phase-encoding, as well as time-frequency analysis investigating neural oscillations. Here, we will start by discussing the results of each analysis separately, and then we will bring them together to answer the proposed questions.

Amplitude-Encoding Analysis

To assess amplitude-encoding, we computed the correlation between the speech envelopes and the EEG evoked responses in the 1-40 Hz range. Mean correlation results ranged from 0.094 to 0.217 for **English**, from 0.082 to 0.149 for **French**, and from 0.068 to 0.130 for **Spanish**. We validated our results by verifying if their distributions were significantly different from zero. We found that correlation distributions from all channels in the English and French condition were significantly higher than zero. However, for Spanish only channels F3 and Fz were significantly higher than zero. We also submitted the correlation coefficients to a permutation test because it represents a more stringent way to evaluate the significance of our results. With this test we were able to show that correlations were significantly higher than chance for all channels when processing English, and for some channels (F7, F3, Fz) when processing Spanish, but not significantly higher than chance when processing French. The permutation results are fairly consistent with the t-test results for English and Spanish. Based on the fact that correlation results for English were significantly higher than zero and chance for all channels, we establish that there is amplitude-encoding for this language. Moreover, based on the fact that correlation results for French and Spanish were not significantly higher than both zero and chance for most channels (10/10 channels for French, and 8/10 channels for Spanish), we establish that there is no amplitude-encoding for these languages. Therefore, we conclude that the brain of 6-month-olds is able to track the

amplitude of the speech envelope of the rhythmically different unfamiliar language (English) but not of the native language (French) or the rhythmically similar unfamiliar language (Spanish). This amplitude-encoding difference across languages is in line with the marginal language effect found in the linear mixed effects model.

When comparing the amplitude-encoding results of 6-month-old infants to those of newborns, we found tracking being enhanced in the older population for English, and no relevant differences for French and Spanish. The higher correlations found for English in the 6-month-old participants (compared to birth) are in line with the language differences found at 6 months, exhibiting better tracking for English than for French and Spanish. Even though the French and Spanish correlations at 6 months were not significantly lower than those at birth, we believe (from a qualitative observation of the comparisons) that they do tend to be lower, and that the high individual variability present in the distributions could have prevented this difference from being significant.

Our amplitude-encoding results at 6 months are of similar magnitude to those observed at birth, and it is only when comparing the correlation distributions from both groups that the differences are found. Therefore, when comparing our amplitude-encoding results at 6 months to the adult results found in the literature the same observations made for the newborn results apply: (1) Our infant results are lower than those found for older children ($r = 0.37$, Abrams et al., 2008) probably due to the less structured nature of the infant EEG (exhibiting less peaks than in older children), and the lower signal to noise ratio present in our data (with only 50 repetitions compared to 1,000 in the study with older children). (2) Our infant results are of similar magnitude to those observed in adults for single trial presentation when tested with ECoG ($r = 0.05$-0.20, Kubanek et al., 2013). However, based on the statistical validation of our infant results, we know that these similarities with the mean correlations found in the literature are not enough to establish envelope tracking.

As it was mentioned before in this thesis, there are no infant studies evaluating speech envelope tracking by means of amplitude-encoding (to our knowledge). There is however one study investigating envelope tracking in 7-month-old English learning infants by means of regression models used to predict the EEG signal (Kalashnikova et al., 2018). In that study they assessed envelope tracking to the native language (English) when listening to infant-directed speech (IDS) and adult-directed speech (ADS). They found no tracking for ADS, and some tracking for IDS. This reference thus gave us some insight into our results: same as Kalashnikova et al. (2018), we did not find envelope tracking for the native language, which

in our study was tested with mild infant-directed speech. Taking into account that we did find envelope tracking to the native language at birth (using the same stimuli than for the 6-month-olds) we hypothesize that speech envelope tracking of the native language has a U-shaped trajectory in development: it is present at birth, not present around 6 and 7 months of age, and then reappears later in childhood (shown in 9 to 13 year-olds, Abrams et al., 2008). To further explore this hypothesis, it will be necessary to carry out more experiments investigating infants acquiring languages that are prosodically different from French, and testing populations at other ages during the first months of life.

We did find however, envelope tracking to the rhythmically different unfamiliar language (English), a condition that has not been tested in the infant literature. We suggest two possible explanations for this outcome. (1) One possible reason could be the acoustic shape of the tested English sentences. English has more variable syllable structure than French and Spanish, producing a less regular amplitude envelope, with more and more irregularly distributed peaks. The presence of these salient acoustic events may facilitate encoding the speech envelope in English. This explanation rules out the effect of language familiarity, meaning that the acquired experience with the native language (French) does not improve the tracking of the speech envelope, and attributes the language differences to the acoustic information of the stimuli, and therefore needs to be further investigated by creating stimuli that systematically vary for these acoustic differences. (2) A second explanation could be due to attunement to the native language. We suggest that the amplitude of the neural response decreases with experience and with the increasingly efficient processing of the native language. This is in line with the hypothesis that starting at around 4 months infants might start to rely more on the phoneme repertoire of their native language (Molnar et al., 2013), so the tracking of the speech envelope subsides. These suggestions are supported by the dip in activation found at 6 months in the theta band (Fig. 33). The processing of a rhythmically different unfamiliar language (English) triggers then a stronger neural response possibly related to stronger processing effort. Infants could be trying to impose the rhythm of their native language (French, syllable-timed) to the rhythmically unfamiliar one (English, stress-timed) because they cannot process it automatically, resulting then in effortful processing that is reflected in better envelope tracking. This is in line with an adult study showing enhanced speech envelope tracking when listening to sentences in a second language (a familiar non-native language) compared to the native language (Song & Iverson, 2018). This was

attributed to increased listening effort leading to increased attention, which in turn led to enhanced neural tracking.

As a final note, we want to point out the importance of performing stringent statistical tests, like the permutation test performed here, to establish a phenomenon, that otherwise could have not been correctly identified based only on more traditional tests, which many existing studies in the literature used (comparison to zero, and linear mixed models).

Phase-Encoding Analysis

To assess phase-encoding, we computed the phase-coherence between the speech envelopes and the EEG evoked responses in the 3-6 Hz range (syllabic rate). Mean phase-coherence results ranged from 0.276 to 0.336 for **English**, from 0.253 to 0.351 for **French**, and from 0.243 to 0.342 for **Spanish**. We validated our results with two statistical tests: (1) comparison to zero, and (2) a permutation test. Additionally, we investigated whether language or channel had an effect on these results. Based on the tests' outcomes, we can conclude that 6-month-olds have the ability to encode the phase of the speech envelope, and that they do it equally well for familiar and unfamiliar languages.

When comparing the phase-encoding results of 6-month-old infants to those of newborns, we found that the effect of age on the ability to encode the phase of the speech envelope was not very relevant for the three languages (only present in 1 or 2 channels). This is in turn reflected by the fact that 6-month-old infants continue to be able to track the phase of the speech envelope for the three languages (familiar and unfamiliar alike), suggesting that this ability is not shaped by language experience. Our phase-encoding results are therefore of similar magnitude to those observed in newborns (Chapter 2) and in adults (Pefkou et al., 2017). This suggests that phase-encoding does not follow the same U-shaped trajectory that was observed in amplitude-encoding.

Amplitude Encoding vs. Phase-Encoding

Different results were observed for envelope tracking by means of amplitude-encoding and phase-encoding: a U-shaped trajectory was observed for amplitude-encoding (with no tracking of the native language at 6 months), while phase-encoding does not exhibit this behaviour (it is present at birth, at 6 months, and in adults). Could this mean that amplitude-encoding and phase-encoding represent two distinct neurophysiological mechanisms of

oscillatory synchronisation (Meyer, 2018)? While both mechanisms encode features of the speech envelope, it is possible that they play different roles in speech processing. Amplitude-encoding (Nourski et al., 2009) as well as phase-encoding (Luo & Poeppel, 2007; Peelle et al., 2013) have been related to speech comprehension in the literature when comparing intelligible and unintelligible speech (compressed speech, noise degraded speech), showing enhanced envelope tracking for the intelligible condition. Some studies however, have questioned this tracking-comprehension relationship by showing similar phase-encoding responses for intelligible and unintelligible speech (time-reversed speech, Howard & Poeppel, 2010; time-compressed speech, Pefkou et al., 2017). These results thus call into question whether speech comprehension and envelope tracking are causally linked. Our newborn results showing amplitude-encoding and phase-encoding of familiar and unfamiliar languages support the hypothesis that both mechanisms are present in the absence of comprehension. Additionally, our results showing that 6-month-olds encode the phase of the speech envelope for familiar and unfamiliar languages equally well suggest that phase-encoding is less modulated by experience and comprehension than amplitude-encoding, as also suggested by the fact that previous studies showing envelope tracking of unintelligible speech assessed it by means of phase-encoding (Howard & Poeppel, 2010; Pefkou et al., 2017). This supports the idea that the two mechanisms play different roles in speech processing. As it was mentioned in the *General Introduction* (Chapter 1), studies evaluating phase-encoding have found it to take place mainly in the lower frequencies (delta and theta bands) (Luo & Poeppel, 2007; Howard & Poeppel, 2010; Peelle et al., 2013; Golumbic et al., 2013), while studies investigating amplitude-encoding have found it in the higher frequencies (usually including the gamma band) (Abrams et al., 2008; Nourski et al., 2009; Kubanek et al., 2013). We suggest one possible explanation for the different relationships that amplitude-encoding and phase-encoding seem to have with comprehension: it is possible that phase-encoding (present in the low frequencies) precedes and modulates amplitude-encoding (present in the higher frequencies), by means of the nesting relationship that exists between low-frequencies and high-frequencies (i.e. the phase of the low frequencies modulates the amplitude of the high frequencies, Giraud & Poeppel, 2012). If phase-encoding represents an early stage during speech processing, it is not surprising that it is present for a wider range of auditory stimuli (intelligible and unintelligible). Once phase-encoding takes place, and the feasibility of decoding the speech signal (intelligibility) is determined, this could influence whether the lower-frequencies modulate the higher-frequencies for further processing to take place (amplitude-encoding and perhaps comprehension).

Time-Frequency Analysis

Besides envelope tracking, which investigated the 1-40 Hz and 3-6 Hz ranges, we also investigated how neural oscillations in a broader frequency range (1-100 Hz) behave during speech processing. To do so, we transformed the EEG responses into time-frequency representations (using Morlet wavelets) to obtain speech evoked neural oscillations. Table 9 summarizes the results obtained from the various analyses that we conducted on our time-frequency responses.

In analysis 1 we found that the amplitude of the neural oscillations in the delta band was higher than all the other frequency bands. This was also observed in the newborn data (Chapter 2), and corresponds to the 1/f power spectrum normally exhibited by brain activity measured with different recording systems (EEG, MEG, fMRI) having stronger power in the slow frequencies than in the high frequencies (He, 2014). Additionally, we found two more effects that relate to each other: we found that the theta band was significantly lower than all the other bands; and that the English amplitudes were significantly lower than those for French and Spanish. From analysis 2 we observed that a dip in the activity of the theta band is significantly lower for English (Fig. 32), and relates to the rhythmic discrimination that takes place between English on one side, and French and Spanish on the other. We will refer to this frequency dip from now on as theta suppression, and we will further discuss it later in this section.

In analysis 2 we investigated how neural oscillations differed across the three tested languages over the time course of stimulation. We found the strongest effect to be present in the delta, theta and high-gamma bands (Fig. 32a). To further explore the origin of these differences, we compared the time-frequency responses of each pair of languages, and found that the theta and high-gamma effects were mainly due to the differences between the rhythmically different languages (English vs French, and English vs Spanish), and that some differences in the high-gamma band also reflected within-rhythmic class discrimination (French vs. Spanish). Neural oscillations in the theta band are significantly higher for French and Spanish compared to English, and no differences are found between French and Spanish. These differences likely reflect the rhythmic discrimination that is present since birth, and that we also observed in our newborn data (Chapter 2). As it was already mentioned for newborns, it makes sense to find rhythmic discrimination being reflected in the theta oscillations, as the theta band (4-8 Hz) relates to the syllabic rate (4-5 Hz), and the syllabic rate correlates with rhythm (Varnet et al., 2017). One thing to note here is that these infants can already

discriminate the rhythmically similar languages (French vs. Spanish), and this discrimination was not reflected at all in the theta band, suggesting that the theta oscillations process similar rhythms in a very similar way. If this were the case for other languages and other rhythmic classes, then theta oscillations could be used to perform rhythmic classification. This will need to be further investigated. In this analysis we also found gamma oscillations to be related to rhythmic discrimination. Language effect is stronger in the high-gamma band than the low-gamma band, however, differences in the low-gamma band (similarly to those in the theta band) are mainly present in the between-rhythmic classes comparisons (English vs. French, and English vs. Spanish) and not so much in the within-rhythmic class comparison (French vs. Spanish). This is similar to what we found for newborns (Chapter 2), suggesting that activity in the low-gamma band also relates to rhythmic processing (treating similar rhythms in a very similar way), and that neither theta nor low-gamma activity are related to language familiarity (as French was not processed differently). For the high-gamma oscillations (50-100 Hz) we do see some new oscillatory behaviours that were not observed in the newborn data: high-gamma activity reflects both between-rhythmic class and within-rhythmic discrimination. High-gamma oscillations show differences related to rhythmic discrimination, having lower activation for the rhythmically different unfamiliar language (English) than for the rhythmically familiar ones (French and Spanish). Additionally, high-gamma oscillations also reflect the discrimination of rhythmically similar languages, as they have higher activation for the native language (French) than for the rhythmically similar unfamiliar one (Spanish). It is important to note that within-rhythmic class discrimination takes place only if one of the languages is familiar to the infant (Nazzi et al., 2000; Nazzi & Ramus, 2003), which is why these differences are a reflection of both language familiarity and discrimination. This higher activation of the high-gamma oscillations when listening to the native language was also shown for 6-month-olds by Peña et al. (2010), and for adults by Peña & Melloni (2012). We also found a strong cross-linguistic effect in the delta band. After the language pairwise comparisons, this effect was only present in the comparisons including Spanish. At the moment, we do not understand this effect. Delta corresponds to the processing of phrasal units (Giraud & Poeppel, 2012), and it is possible that this effect is due to specific features present in the tested Spanish sentences. To confirm the abovementioned findings, we will re-run this cross-linguistic analysis with a permutation test to control for multiple comparisons.

In analysis 3 we investigated how the neural oscillations related to speech processing change after the first 6 months of life. Here we identified two main developmental effects: (1) differences in the lower frequency bands (delta, theta, alpha) are present for the three languages, so they do not derive from language familiarity, but likely reflect brain maturation, general experience with language, and/or more general cognitive development. (2) Differences in the gamma band (low and high) are mainly present for the native language (French) and the rhythmically similar unfamiliar language (Spanish), likely reflecting language familiarity. The changes in the lower frequencies exhibit two patterns: an initial increase of activity in the three frequency bands (delta, theta, alpha) that after approximately 500 ms diminishes for theta and alpha, but remains for delta. This is followed by a dip in the theta activity, which we previously referred to as the theta suppression. One possible explanation for this behaviour is a short increase of attention right after sentence onset (reflected by the increase in alpha activation), paralleled by the processing of the sentence's prosody (reflected by the increase in delta and theta). The delta activation remains during the time course of sentence presentation (as also observed in Fig. 30) suggesting the continuous processing of phrasal units. The theta suppression suggests that after the rhythm of the language has been processed at the beginning of the sentence (initial peak in theta activity) the focus shifts towards smaller speech units (likely reflected by activity in the gamma band) as at this age processing of individual phonemes becomes important, as it corresponds with the beginning of word and grammar learning. In other words, we argue that syllable suppression takes place in order to be able to better represent phonemes. The other main developmental change observed here is the one reflected in the gamma-band. Differences in the gamma band are mainly present in the native language (French) and the rhythmically similar unfamiliar language (Spanish) by exhibiting higher activation of these neural oscillations at 6 months than at birth. The changes in the gamma band are not present for English, therefore we hypothesize that they are due to experience with the native language (French). We want to point out that even though the increase in gamma activity is also present for Spanish (which is an unfamiliar language), gamma oscillations are stronger when processing French, as it was also shown in analysis 2 (Fig. 32d). Therefore, we believe that the developmental increase in gamma activity shown for Spanish is a reflection of rhythmic familiarity. This increase in gamma activity is consistent with what was previously mentioned about theta suppression occurring in order to shift the focus towards the processing of phonemic units. To account for the different sample sizes in the two age groups (47 participants at birth, and 23 participants at 6 months), we run additional analyses on matched

size subsamples and found the same main effects described above. To confirm the abovementioned developmental differences, we will re-run this analysis with a permutation test to control for multiple comparisons.

Research Questions & Answers

In sum, let us return to the questions that motivated this study:

(i) How do neural oscillations relate to between-rhythmic class discrimination (French vs English, Spanish vs English) at 6 months? The discrimination of languages from different rhythmic classes was represented by neural oscillations in the theta and gamma bands, were neural activity was stronger for the rhythmically familiar languages (French and Spanish) than for the rhythmically unfamiliar one (English). This is consistent with what was found in newborns in which case rhythmic discrimination was represented by the theta and low-gamma bands.

(ii) How do neural oscillations relate to within-rhythmic class discrimination (French vs Spanish) at 6 months? The discrimination of languages from the same rhythmic class was represented by neural oscillations in the high-gamma band, were neural activity was stronger for the native language (French) than for the rhythmically similar unfamiliar one (Spanish). This is consistent with what has been shown in the literature (Peña et. al., 2010; Peña & Melloni, 2012).

(iii) How do neural oscillations change after the first 6 months of life? When comparing the speech processing capacities of 6-month-olds to those of newborns two main developmental changes need to be taken into account: maturation and speech experience. We argue that due to maturation the neural oscillations operating in the lower frequency bands (delta, theta, alpha) developed to initially process the prosody of the incoming stimuli (initial peak in activation) and to then suppress their activity (theta suppression) to allow other frequency ranges to process other speech units. This is in line with the hierarchical relation that has been observed in adult neural oscillations during speech processing (Giraud & Poeppel, 2012), and we will further investigate it in future analyses (see *Future Directions*, Chapter 5). We also suggest that due to speech experience neural oscillations in the gamma band became attuned to the characteristics of the native language, as reflected by the stronger neural activation found when processing the familiar language (French), and the rhythmically similar unfamiliar one (Spanish), and is likely related to the onset of word and grammar learning.

(iv) How do language experience and age affect the synchronization of neural oscillations to the speech envelope (speech envelope tracking)? Here we assessed envelope tracking using two measures: amplitude-encoding and phase-encoding. The tracking of the envelope's phase did not seem to change with age nor with language experience. At 6 months of age infants continue to track the phase of the speech envelope in familiar and unfamiliar languages equally well. However, this is not the case for the amplitude of the speech envelope. At 6 months of age infants seem to go through a developmental period in which the encoding of the speech amplitude is suppressed, and it is only regained under certain circumstances: in the case of a very attractive speech input (as shown for infant directed speech by Kalashnikova et al. (2018), or an unfamiliar language (as show by our results for English).

CHAPTER 4: Predicting language outcome

4.1 INTRODUCTION

Even though most children acquire language effortlessly, not all do. Around 7% of kindergarten children (5-6 years) (Tomblin et al., 1997) are identified as having specific language impairment (SLI), a disorder characterized by the difficulty to understand and produce spoken language in the absence of other cognitive deficits. Another 5% to 17% of school children suffer from dyslexia (Shaywitz, 1998), a specific deficit in reading acquisition not attributable to low IQ, poor education or neurological damage (Ramus & Ahissar, 2012). If untreated, these disorders can have an impact on many aspects of the child's life (social, behavioural, academic), which can persist until adulthood. It is, therefore, fundamental to have diagnostic tools that could allow the early detection of such language delays. Nowadays language impairments are difficult to diagnose before 3-4 years of age (Cristia et al., 2014), because diagnosis relies on behavioural criteria that are difficult to obtain early in life. However, since these disorders are believed to have a biological / genetic component, i.e. an alteration in the nervous system, it is believed that manifestations may be present in the brain from early on, so it is plausible that they could be detected at an earlier age, than currently done, by means of neural measures.

A neural measure can be a predictor of language development if it can explain a percentage of the variation observed in language outcome, even within the healthy population. If the percentage explained is very high, then the neural measure represents a very strong predictor that could have some clinical relevance. However, if the percentage explained is small, but significant, the neural measure is still considered a predictor (statistically speaking) but it would not have clinical relevance. The work presented in this chapter focuses on identifying language predictors that could explain later vocabulary measures, but given our methodology, we cannot aim to identify clinically relevant predictors.

To use a neural measure as a potential predictor of later language capabilities, it is important to look at the individual variability present in the measure, as it could be informative in describing the developmental trajectory of each individual. To explore whether neural oscillations could be used as a predictor of language outcome, we followed up the infants tested in the EEG study at ages 0 and/or 6 months during their first 2 years of life to assess the growth of their vocabulary. Specifically, we collected information about their receptive and expressive vocabulary at 12 and 18 months using the MacArthur-Bates Communicative Developmental Inventory (CDI) questionnaires. In this chapter, we investigate the linkage between the early neural responses and later language development.

The MacArthur-Bates Communicative Developmental Inventory is a questionnaire frequently used in studies investigating the relationship between early speech perception abilities and later language outcome. This questionnaire was designed for parents to report information about their children's' developing language abilities. It contains checklists to collect information about vocabulary comprehension, production, gestures and grammar, according to the age that is being targeted. For infants between 8-16 months old the *Words and Gestures CDI* provides checklists to measure vocabulary comprehension, vocabulary production, and gesture production. For older children, between 16-30 months, the *Words and Sentences CDI* measures language production (Fenson et al., 1994). The use of these questionnaires to assess language and communication skills has been shown to be reliable and valid from 8 to 30 months of age (Fenson et al., 1994). With this longitudinal design we aimed to find early neural measures of speech perception that could predict later language outcomes.

4.2 MATERIALS AND METHODS

4.2.1 Participants

The parents of the infants who participated in at least one of the EEG experiments (at birth or at 6 months) were requested to fill up vocabulary questionnaires when their children turned 12 and 18 months. As it is often the case in longitudinal studies, some of the participants did not contribute with all four assessments. Table 10 presents the list of participants and the assessments that they contributed longitudinally. A total of 47 participants contributed neural data at birth, while 25 participants were tested at 6 months (12 of these participants contributed both neural assessments). Additionally, vocabulary data was collected from 36 participants at 12 months, and from 35 participants at 18 months (27 of these participants contributed both vocabulary assessments). Only 8 participants contributed all four assessments.

4.2.2 Procedure

We contacted by post the families that had agreed to be part of the longitudinal study, and requested them to fill up a vocabulary questionnaire when their child turned 12 and 18 months. In each case we asked them to return the questionnaire before their child turned 13 and 19 months respectively, to ensure that the measurement would not exceed these age limits.

Baby	EEG (birth)	EEG (6 months)	CDI (12 months)	CDI (18 months)	Baby	EEG (birth)	EEG (6 months)	CDI (12 months)	CDI (18 months)
1	X				31	X		X	X
2	X	X	X	X	32	X		X	
3	X	X		X	33	X	X	X	X
4	X	X	X	X	34	X			X
5	X		X	X	35		X	X	X
6	X		X	X	36	X			
7	X		X	X	37	X			X
8	X				38	X			X
9	X	X	X	X	39	X		X	X
10	X	X	X	X	40	X			
11	X			X	41	X		X	X
12	X	X	X	X	42	X		X	X
13	X	X	X	X	43	X		X	X
14	X		X	X	44	X	X	X	X
15	X			X	45	X			X
16	X				46	X			
17	X	X			47	X			
18	X				48	X			
19	X		X		49	X			
20	X		X		50	X		X	X
21		X	X	X	51	X	X	X	X
22		X	X	X	52		X		
23	X				53		X	X	
24		X			54		X		
25	X	X	X		55		X	X	
26	X				56		X	X	X
27	X		X	X	57		X	X	X
28	X		X	X	58		X	X	
29	X		X		59		X	X	X
30	X		X	X	60		X		

Table 10. Summary of data points collected for each participant longitudinally

To obtain a measure of language development we used the French version of the MacArthur-Bates Communicative Developmental Inventory (CDI) questionnaires (Kern, 2007). To make it easier for parents to complete the questionnaires and to improve our chances of obtaining a response, we provided them with the short version of the CDI, which is only one page long. The short version CDI has been shown to be as reliable as the original version for the English CDI (Floccia et al., 2018). For the measurement at 12 months we used the *Words and Gestures* CDI, which inquires about the child's babbling skills, provides a list of 83 words for parents to indicate whether the child understands them and spontaneously produces them, and a list of 25 gestures for them to indicate if the child makes them (e.g. shake the head to say no). For the measurement at 18 months we used the *Words and Sentences* CDI, which provides a list of 97 words for parents to indicate whether the child understands them and spontaneously produces them, and inquires whether the child has started to combine words together.

4.2.3 Analysis of Longitudinal Data

The first step in the analysis of the longitudinal data was to define which measures to include from each age in a hypothesis-driven, principled way. This was particularly important to avoid spurious results, as our dataset is of high dimensionality.

For the neural data recorded at birth and at 6 months we had various measures to consider as possible predictors: envelope tracking measures (correlation and phase-coherence), and neural oscillations measures (at different frequency bands). We decided to choose measures that were modulated by experience/familiarity, as these are more likely to be informative about native language development (Kuhl et al., 2005). Consequently, based on our results that envelope tracking is not modulated by language experience, we decided to not use these measures (correlations and phase-coherence) in the longitudinal analysis. By contrast, certain frequency bands of the neural oscillations were modulated by language experience, in particular, theta and low-gamma at birth as well as theta and gamma (low and high) at 6 months. We thus focused on these frequency bands elicited by the native language (French). From the CDI data collected at 12 and 18 months, we selected the number of *words understood* as a measure of receptive vocabulary, and the number of *words produced* as a measure of expressive vocabulary.

As a first analysis, we evaluated the reliability of our vocabulary data and tracked infants' vocabulary growth. To this effect, we computed the correlation between measures at 12 and 18 months for the receptive and the productive vocabulary scores. The sample size for these comparisons was 28 participants.

As a second analysis, we directly addressed the question of whether neural oscillations may predict later vocabulary growth. We computed the Pearson correlation between the neural measures (theta and gamma at birth and at 6 months) and the vocabulary measures (words understood and words produced at 12 and 18 months) to explore if there was a linkage between them. For the neural data, we used the mean spectral amplitudes in the time-frequency windows in which significant activation was found. For the theta band (4-8 Hz) at birth, this was between 1200 and 1800 ms (Fig. 18 in Chapter 2); at 6 months 400 – 1900 ms (Fig. 30 and Fig. 32b in Chapter 3). For the low-gamma (31-50 Hz) and high-gamma (51-100 Hz) bands, we included the whole time range, because significant differences in these bands were not restricted to specific time windows.

We hypothesize that the theta suppression and the gamma increase observed at 6 months are related to the beginning of the attunement to the native language, and word/grammar

learning. Therefore, we expect these measures to correlate with the CDI data: negatively for the theta suppression (6 months) and positively for the gamma increase.

4.3 RESULTS

4.3.1 Vocabulary Growth

We assessed the trajectory of the vocabulary measures by computing the correlation between data at 12 and 18 months. The trajectories for word comprehension and word production are displayed in Fig. 36a and Fig. 36b, respectively. They exhibit a vocabulary growth that is consistent across participants, and significantly correlated in time: r = 0.755 (p-value < 0.01) for word comprehension at 12 and 18 months, and r = 0.660 (p-value < 0.01) for word production at 12 and 18 months. These results confirm that CDI questionnaires provided reliable measures of language growth in our sample.

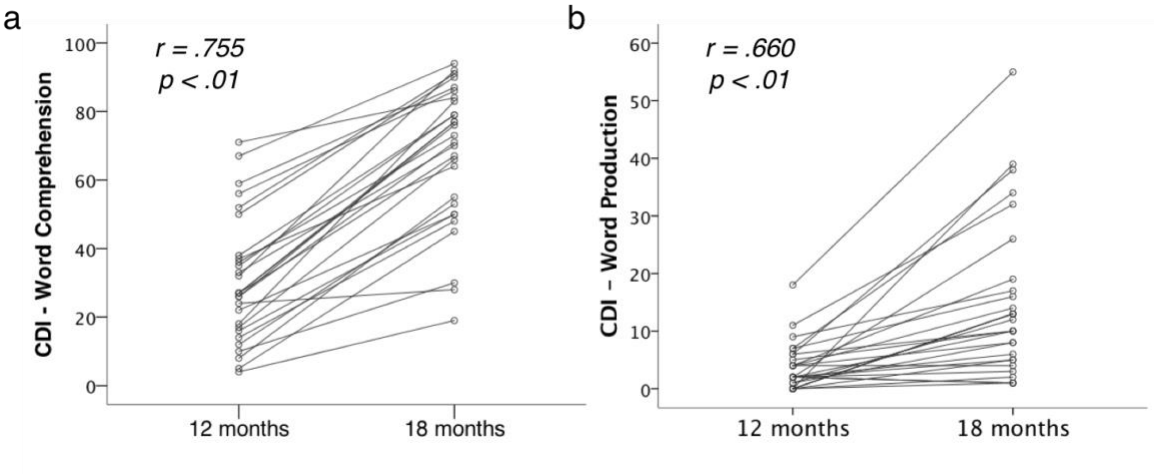


Figure 36. Trajectory of vocabulary measures collected with CDI questionnaires. **a)** The number of words understood at 12 months is correlated with the number of words understood at 18 months (r = 0.755, p-value < 0.01). **b)** The number of words produced at 12 months is correlated to the number of words produced at 18 months (r = 0.660, p-value < 0.01).

4.3.2 Predicting Language Development

After confirming the reliability of the vocabulary data, we explored the linkage between the neural measures and the vocabulary measures by computing the Pearson correlation between them. Table 11 and Table 12 present the results of these comparisons for each age group (newborns and 6-month-olds). To correct for the multiple comparisons computed on each dependent variable, i.e. to adjust the family-wise error rate for each comparison (n=4), we used the Bonferroni correction to adjust the original alpha value (α=0.05). This resulted in the

adjusted alpha value (α = 0.0125), which was used to evaluate the obtained results. As it can be observed in Table 11 and Table 12, only one correlation between the neural measures and the vocabulary measures was significant: theta spectral amplitude at birth correlates with the number of words produced at 12 months (r = 0.494, p-value = 0.009, alpha = 0.0125). Figure 37 displays the individual measures that account for this significant correlation.

Neural measure	Vocabulary Measure	Correlation Coefficient	p-value	Sample Size
Theta (0m)	Comprehension (12m)	0.135	0.503	27
Theta (0m)	Production (12m)	0.494	**0.009**	27
Theta (0m)	Comprehension (18m)	0.054	0.782	29
Theta (0m)	Production (18m)	-0.100	0.606	29
Low-Gamma (0m)	Comprehension (12m)	-0.017	0.934	27
Low-Gamma (0m)	Production (12m)	-0.079	0.696	27
Low-Gamma (0m)	Comprehension (18m)	-0.002	0.991	29
Low-Gamma (0m)	Production (18m)	-0.124	0.521	29

Table 11. Correlation coefficients between neural measures at birth and vocabulary measures at 12 and 18 months. The alpha value for this test is α = 0.0125.

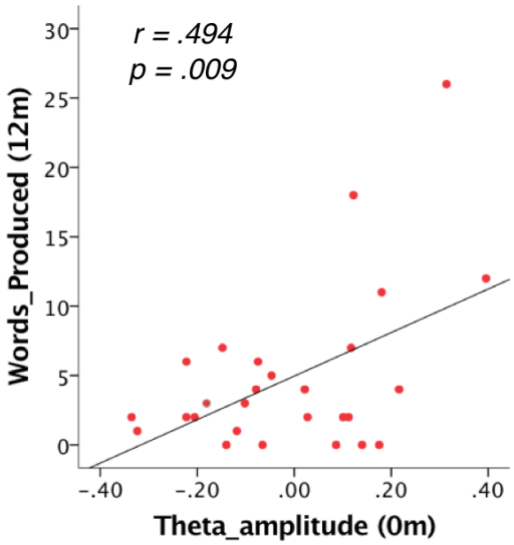


Figure 37. Correlation between theta spectral amplitude at birth, and the number of words produced at 12 months (r = 0.494, p-value = 0.009). The alpha value for this test is α = 0.0125.

Neural measure	Vocabulary Measure	Correlation Coefficient	p-value	Sample Size
Theta (6m)	Comprehension (12m)	-0.154	0.555	17
Theta (6m)	Production (12m)	0.334	0.190	17
Theta (6m)	Comprehension (18m)	-0.314	0.274	14
Theta (6m)	Production (18m)	-0.043	0.884	14
Low-Gamma (6m)	Comprehension (12m)	0.506	**0.038**	17
Low-Gamma (6m)	Production (12m)	-0.108	0.681	17
Low-Gamma (6m)	Comprehension (18m)	0.309	0.282	14
Low-Gamma (6m)	Production (18m)	0.021	0.944	14
High-Gamma (6m)	Comprehension (12m)	0.467	**0.059**	17
High-Gamma (6m)	Production (12m)	-0.142	0.587	17
High-Gamma (6m)	Comprehension (18m)	0.373	0.189	14
High-Gamma (6m)	Production (18m)	0.224	0.442	14

Table 12. Correlations between neural measures at 6 months and vocabulary measures at 12 and 18 months. The alpha value for this test is $\alpha = 0.0125$.

We note, however, that two correlations, highlighted in bold in the Table 12, were significant or marginally significant before correction. These included low and high gamma spectral amplitudes at 6 months predicting comprehension scores at 12 months. Figure 38 (a,b) displays the individual measures that account for these marginal correlations.

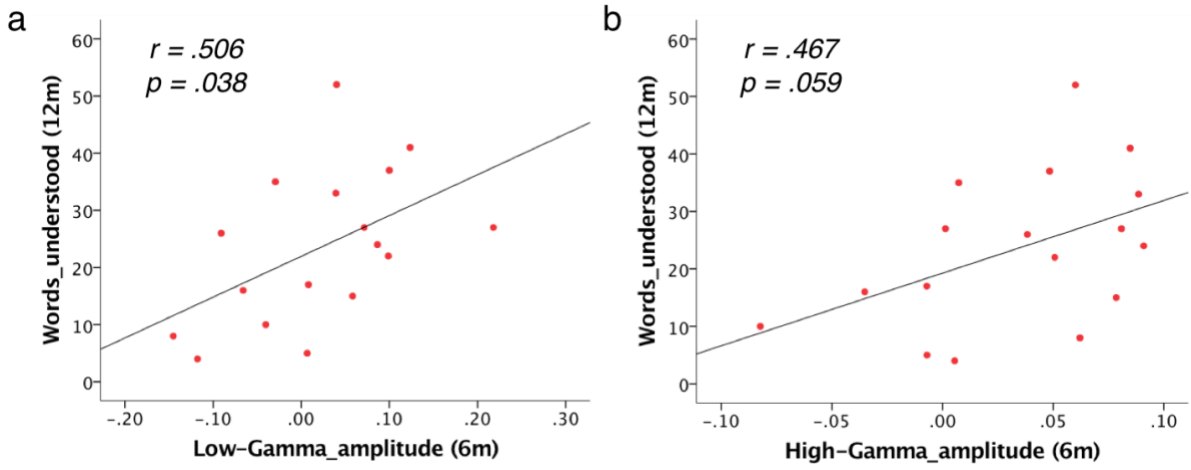


Figure 38. Marginally significant correlations. **a)** Low-gamma at 6 months vs. number of words understood at 12 months (r = 0.506, p-value = 0.038). **b)** High-gamma at 6 months vs. number of words understood at 12 months (r = 0.467, p-value = 0.059). The alpha value for these comparisons is $\alpha = 0.0125$.

4.4 DISCUSSION

In this chapter we investigated whether neural oscillations could be used as a predictor of language outcome. We selected theta and low-gamma oscillations at birth, and theta and gamma (low and high) oscillations at 6 months as plausible candidates for predictors due to modulation by experience with the native language. To assess the linkage between these possible predictors and later language outcome, we computed the correlation between the mean spectral amplitude at each frequency band (in a given time-window) and two measures of vocabulary size at 12 and 18 months: number of words understood and number of words produced.

Our results show that one of the chosen measures of neural oscillations relates to later language measures: theta spectral amplitude at birth correlates with the number of words produced at 12 months. The predictive power of the theta activity at birth could be due to a strong theta activation being related to a good encoding of syllables, which may be due to inherent abilities or even prenatal experience. Since French is a syllable-timed language, and theta is argued to be crucial for speech perception in general (Ghitza, 2011), this stronger theta activation could indicate stronger speech perception abilities, leading to better production. To further investigate this, we will explore (in future analyses) if theta activity at birth elicited by the unfamiliar languages also correlates with production at 12 months. This would test whether the effect is based on prenatal experience (if only the correlation with French is significant) or general abilities (if the other languages are also significant).

Other two correlations were significant or marginally significant before the multiple comparisons correction. Results suggest that gamma band oscillations at 6 months may be predictors of vocabulary size at 12 months. However, this relationship is not significant, and we believe that this is due to the small sample size (17 participants). We interpret these marginal results as an indication of a possible relationship that is worth exploring further on a larger sample, yielding more statistical power.

Importantly for our developmental perspective, the significant or marginally significant predictors were exactly those frequency bands that our hypotheses predicted to be most strongly modulated by experience at a given age. Thus prenatal experience, lacking high frequency information related to phonemes, was predicted to affect slower oscillations, whereas the full-band postnatal experience was predicted to shape the faster bands responsible for phonemic processing by 6 months of age. This is exactly what we found: the theta band is a good predictor at birth, while the gamma band is at 6 months.

A second reason that might have contributed to the weakness of the results obtained is that our study is the first to explore neural oscillations as predictors of later language outcome. Therefore, there was no existing evidence as to what measures would be most appropriate to use for the prediction. We decided to use the frequency bands in the time-frequency windows that showed significant modulation by the native language. However, individual variability in these measures may be too high and some other method of selecting the relevant oscillatory measures may be more appropriate.

A third factor that could have played an important role is that of data quality. To use a neural measure as a predictor of language outcome, it is fundamental that the neural measure be reliable. The variability elicited by the different possible behavioural states of participants (calm vs. fussy etc.) might be negligible when looking at the group results, but impactful when looking at individual results. For this reason, it would be informative to check for the test-retest reliability of the neural measures to guaranty that good quality, reliable and replicable neural predictors are used.

In sum, we have found here some indication of a predictive relationship between early neural oscillations and later language outcomes that merits further exploration and confirmation in future studies. These results point towards a developmental scenario in accordance with our theoretical predictions as well as the empirical finds of the time-frequency analyses: the frequency bands that encode the developmentally relevant speech experience (the theta band encoding slower modulations in speech at birth and the gamma band encoding faster modulations in speech at 6 months) are the ones that serve as predictors for later language development.

CHAPTER 5: General Discussion

Neural oscillations have been proposed to underlie speech processing and comprehension (Giraud & Poeppel, 2012; Meyer, 2018). Up to date, the relationship between neural oscillations and speech has mainly been investigated in adults, a population of experienced listeners. We know from animal studies that neural oscillations are not exclusive of humans, and not specific to speech and language. However, it is unknown whether the relationship between neural oscillations and speech as observed in adults is in place due to experience with speech and language acquisition, or whether it is in place from the beginning of life. The present work had two main goals: (1) to investigate the developmental origin of neural oscillations and their relationship to speech processing, and (2) to explore the developmental changes of this relationship during the first 6 months of life.

Below, we will first summarize the obtained results, we will then discuss their implications for the development of neural oscillations and language learning.

5.1 SUMMARY OF RESULTS

5.1.1 Resting State Neural Oscillations

In Chapter 2, we investigated resting state neural oscillations in the newborn brain before and after speech stimulation. At rest, activity in the newborn cerebral cortex is mainly present in the delta oscillations (1-3 Hz) (Fig. 15 in Chapter 2), which is similar to what has been shown for adults (Fig. 1a in Chapter 1, Giraud & Poeppel, 2012). We also found that newborn resting state neural oscillations were modified after speech stimulation took place. After a stimulation of approximately 21 minutes, neural oscillations did not go back to (pre-stimulus) baseline and exhibited higher amplitudes in all the frequency bands of interest (delta, theta, alpha, beta, gamma) during different time-frequency windows.

5.1.2 Neural Oscillations and Speech Processing

In Chapter 2, we investigated the role of neural oscillations during speech processing of familiar (French, heard prenatally) and unfamiliar (Spanish, rhythmically similar to French, and English, rhythmically dissimilar) languages at birth. We found that shortly after birth, neural oscillations in the theta, beta and low-gamma bands already seem to relate to speech processing. More specifically, theta (4-8 Hz) and low-gamma (31-50 Hz) oscillations seem to be modulated by language, as they exhibited stronger activation for the rhythmically familiar languages (French and Spanish) than for the rhythmically unfamiliar one (English) (Fig. 17 in Chapter 2). Beta oscillations (13-30 Hz) on the other hand do not seem to be modulated by

language, and their activation appears to be triggered by all three languages alike (Fig. 15 and Fig. 18 in Chapter 2).

In Chapter 3, we investigated how the relationship between neural oscillations and speech processing changed after the first six months of life. We found that at 6 months, theta (4-8 Hz) and low-gamma (31-50 Hz) oscillations continue to be modulated by language rhythm, exhibiting stronger activation for the rhythmically familiar languages (French and Spanish) than for the rhythmically unfamiliar one (English) (Fig. 32 in Chapter 3). Additionally, we observed the emergence of high-gamma (51-100 Hz) oscillations that were not present at birth. High-gamma oscillation also appeared to be modulated by language familiarity, exhibiting stronger activation for the native language (French) than for the rhythmically similar unfamiliar language (Spanish) and for the rhythmically different unfamiliar one (English) (Fig. 32 in Chapter 3). When assessing developmental changes from birth to 6 months of age we identified two main effects: (1) developmental changes in the delta (1-3 Hz), theta (4-8 Hz) and alpha (9-12 Hz) oscillations were present at 6 months for the three languages and were likely to reflect brain maturation and/or general cognitive development; (2) changes in the low-gamma (31-50 Hz) and high-gamma (51-100 Hz) oscillations were mainly present in the rhythmically familiar languages (French and Spanish) and were likely to reflect experience with the native language (French) (Fig. 33 in Chapter 3). At 6 months, the low frequency changes (delta, theta and alpha) exhibited a biphasic activation pattern, not yet present at birth, with an initial increase in activity (from 0 to 500 ms), followed by a dip in theta spectral amplitude, referred here as theta suppression.

5.1.3 Speech Envelope Tracking

In Chapter 2, we investigated whether the newborn brain is able to track the amplitude and the phase of the speech envelope in familiar (French, heard prenatally) and unfamiliar (Spanish and English) languages. We found that the ability to encode the amplitude and phase of the speech envelope is already in place at birth, and that it is independent of language familiarity.

In Chapter 3, we investigated whether the ability to track the speech envelope would change after experience with the native language was acquired during the first 6 months of life. We found that at 6 months, the phase of the speech envelope continues to be encoded by the infant brain independently of language familiarity. However, the encoding of the amplitude of the speech envelope ceases to take place when processing rhythmically familiar languages (French and Spanish), and is present only for the rhythmically unfamiliar language (English).

5.1.4 Linkage Between Neural Oscillations and Language Outcome

In Chapter 4, we investigated whether measures of neural oscillations could be used as potential predictors of language development. We found evidence that theta activity at birth predicted the number of words produced at 12 months. Additionally, we identified other two suggestive effects: (i) low-gamma activity at 6 months predicted the number of words understood at 12 months, and (ii) high-gamma activity at 6 months predicted the number of words understood at 12 months.

5.2 IMPLICATIONS FOR LANGUAGE LEARNING

The above results show that the neural oscillatory mechanisms found in adults are already operational from birth on, and undergo important changes as a result of language experience. How do these oscillations support speech perception and language development? How are they shaped by experience?

Below, we will discuss the implications of our findings for language acquisition. We note that this is one of the first studies investigating neural oscillations in such young infants. This represents, therefore, an exploratory study. Also, not all analyses could be completed and further ones are planned to confirm the current results.

5.2.1 Resting State Neural Oscillations

Our qualitative evaluation of resting state neural oscillations at birth shows that activity is mainly present in the delta band (1-3 Hz). Even though we did not investigate whether neural oscillations at birth already exhibit the nesting relationship that has been observed in adult humans (during speech processing, Giraud & Poeppel, 2012) and animals (at rest, Lakatos et al., 2005), we hypothesize that the presence of delta oscillations might lay the foundation for this hierarchical organization. Neural oscillations are hierarchically nested in animals and adults: the phase of the delta oscillations modulates the amplitude of theta oscillations, and the phase of theta modulates the amplitude of gamma oscillations (Giraud & Poeppel, 2012; Lakatos et al., 2005). Up to date, it is unknown whether this nested relationship is in place at birth. We hypothesize that the fact that we see slow oscillations (delta) in place at birth at rest suggests that the nested hierarchical organization may arise as a function of brain maturation and experience. This later emergence was observed for gamma oscillations, which were only weakly present at birth (during rest and speech processing), and by the age of 6 months exhibited significantly stronger activation (during speech processing). This hypothesis will be

further investigated in the future by exploring the nesting relationship of neural oscillations at birth and at 6 months (*Future Directions*).

Additionally, when investigating the after-effects of speech stimulation, we found that newborn resting state neural oscillations were modified after speech stimulation took place. We interpreted this as an indication of some short-term learning that takes place while the neural oscillations continue to stay active. This is in line with what was found for 3-month-olds who exhibited higher connectivity during resting state after 3 minutes of speech stimulation (Homae et al., 2011).

5.2.2 Neural Oscillations and Speech Processing

Rhythmic discrimination

An important and well-documented ability of newborn infants is to discriminate rhythmically different languages, even if they never heard them before (Nazzi et al., 1998; Ramus et al., 2000). This ability is believed to underlie early multilingual development (Mehler et al., 1988; Ramus et al., 2000; Byers-Heinlein et al., 2010), specifically young infants' ability to discover that they are receiving dual language input, and native language rhythm remains a key factor in speech perception even in adults (Cutler 1994, 2002). Despite the importance of this ability in speech perception, its neural basis is poorly understood, especially in development. Oscillations are natural candidates to encode language rhythm, given their rhythmic nature. Our findings thus shed light on the development and the neural underpinnings of language rhythm and rhythmic discrimination.

Both at birth and at 6 months, theta (4-8 Hz) and low-gamma (31-50 Hz) oscillations appear to be modulated by the rhythm of the familiar language (French), and they do not behave differently for rhythmically similar familiar and unfamiliar languages (French and Spanish). Activity in these two bands thus seems to underlie the ability infants possess to discriminate their native language from a rhythmically different one. If theta and/or low-gamma oscillations encode speech rhythm, as we have found, it is likely that they are responsible for the discrimination of rhythmically different languages. The rhythmic discrimination tested in our study involved a rhythmic class that was familiar to the participants (French and Spanish are syllable-timed languages). It remains to be tested whether theta and/or gamma oscillations would still relate to the rhythmic discrimination of languages from two unfamiliar rhythmic classes (e.g. a stress-timed languages vs. a mora-timed languages in a French learner). Such test would be informative of whether theta and

gamma bands relate to the general ability of discriminating rhythmically different languages, regardless of familiarity.

Since we only tested infants learning one language, French, it remains to be tested whether the theta and low-gamma oscillations specifically encode one type of rhythm, the syllable-timed one, characteristic of French, or whether these oscillations encode the prenatally heard rhythm, whatever that may be. Further studies investigating newborns exposed prenatally to different languages will allow us to tease these two possibilities apart.

At 6 months we observed the presence of high-gamma (51-100 Hz) oscillations, which were not evidenced at birth. Therefore, high gamma oscillations emerge some time during the first 6 months of life, at least as measured by scalp EEG. High-gamma oscillations appear to be modulated by rhythm, because they exhibit higher activation for an unfamiliar language that is rhythmically similar to the native language than for a rhythmically different unfamiliar language. Furthermore, the high-gamma band also appears to encode linguistic information unique to the native language, as higher activation was observed for French when compared to Spanish (the rhythmically similar language). This high-gamma increase exclusive of the native language was also found by Peña et al. (2010) testing 6-month-olds with a similar paradigm. These findings support the hypotheses that infants rely on the phonemic information of their native language to discriminate it from a rhythmically similar one (Bosch & Sebastian-Galles 1997; Molnar et al., 2013), and that gamma oscillations are involved in the processing of this phonemic information (Giraud & Poeppel, 2012).

From a more general developmental perspective, the increase in gamma power and high-gamma-based within-rhythmic discrimination coincide with the beginning of word learning (Bergelson & Swingley, 2012), a process in which phonemic information plays a crucial role, as phonemes are the minimal distinctive units of words. It thus seems developmentally plausible that (high-)gamma activity supports these developmental processes.

Predictive coding from linguistic knowledge

In adults, (de)activation in the beta band is associated with predicting upcoming linguistic content on the basis of existing linguistic knowledge (Howard & Poeppel, 2010, 2012; Pefkou et al., 2017). While young infants have much less linguistic knowledge than adults, it is interesting to ask whether their existing knowledge, e.g. from prenatal or early postnatal experience, allows them to show this classical neural signature of predicting upcoming linguistic information.

We found that beta oscillations (13-30 Hz) are in place from birth. However they do not appear to be modulated by language familiarity, and instead seem to relate to speech processing in general. However, we also found that beta activation was particularly strong towards the end of the sentence. This may be a first indication that beta oscillations also play a predictive role in newborns. One possibility is that newborns may be able to predict the end of an utterance based on its gross prosodic features. This prediction may be broadly based and not specific to the prosody of the native language. Rather, it may rely on the broader, more universal features of utterance-level / intonation phrase level prosodic contours, such as a decrease in pitch/intensity, and an increase in duration (Nespor & Vogel, 1984; Grieser & Kuhl, 1988). Another possibility is that prediction arises due to sentence repetition. This latter hypothesis will be assessed in future analyses of our dataset. Specifically, we will compare beta band activity in early trials and late trials within the language blocks to test whether repeated exposure might account for the observed final increase.

At 6 months of age, activity in the beta band still appears to not be modulated by language familiarity. Even though at 6 months, we have no quantitative measures of how much beta oscillations differ from resting state oscillations, since 6-month-olds were not tested in the resting state condition, a qualitative observation of the group mean time-frequency responses does not reveal the strong activation that was observed at birth towards the end of the sentences. This suggests that to the extent that beta activity may already signal prediction in infants, this predictive role may not be due to the sentence repetitions, as in that case, it should also be present at 6 months. If it were instead related to prosodic features, then its absence at 6 months could be explained by infants shifting their attention from large prosodic units to smaller phonological units such as phonemes, which is consistent with the theta suppression and the absence of amplitude envelope tracking also observed at this age (more details in the following sections).

5.2.3 Developmental Changes in Neural Oscillations

Developmental changes in the lower frequency oscillations (delta, theta, alpha) were present for familiar and unfamiliar languages alike, and therefore are likely to relate to brain maturation and/or general cognitive development.

One important change is the appearance of an initial amplitude increase in the three low frequency bands. This is likely similar to the N100 component typically observed in ERP studies in the auditory domain and considered, together with the P200 component an early

marker of auditory processing onset, generated in the primary and association auditory cortices (Steinhauer & Connolly, 2008).

Another interesting developmental change concerns the activity in the theta band, which exhibits a dip after the above discussed initial peak. The dip begins at around 500 ms and lasts throughout the sentence. We refer to it as theta suppression. We suggest that the initial peak in the theta band within the N100-like event specifically reflects the processing of rhythmic/syllabic information, which, once processed, is suppressed to give way to the processing of other linguistics features such as phonemic information (as suggested by the increase in gamma activity). The suggestion that rhythm is processed in the first 500 ms is also supported by the pairwise language comparisons (Fig. 32 in Chapter 3) showing that rhythmic discrimination takes place at around 500 ms.

Importantly, theta suppression appears at 6 months, exactly the age at which infants start to attune to their native language, and start learning word forms. Both of these processes require increasingly stable phonemic representations. Suppressing syllabic information would allow infants to focus on and better represent individual phonemes.

Interestingly, in adults, a suppression response with a similar latency is often observed in the beta band during speech processing (Howard & Poeppel, 2012). Since infants' electrophysiological responses are typically slower than those of adults (De Haan 2013), it is not impossible that the infant equivalent of adult beta responses manifest themselves in a somewhat slower frequency range, e.g. in the frequencies corresponding to the theta band, although we have not observed such a downward shift in our data more generally. Further analysis will be required to determine if the adult beta suppression is linked to the theta suppression observed in our study.

Developmental changes were also observed in the low-gamma (31-50 Hz) and high-gamma (51-100 Hz) bands and were mainly present in the rhythmically familiar languages (French and Spanish), likely reflecting the acquired experience with the native language (French), as discussed above.

5.2.4 Speech Envelope Tracking

In addition to activation in the different frequency bands, we also examined more specifically whether infants track the speech envelope, a mechanism argued to be crucial for speech comprehension and language processing (Ahissar et al., 2001; Nourski et al., 2009; Luo & Poeppel, 2007; Peelle et al., 2013).

At birth we found amplitude and phase envelope tracking to take place for familiar and unfamiliar languages equally well. These findings support the hypothesis that envelope tracking takes place even in the absence of comprehension, and suggest that it represents a basic auditory ability that does not require knowledge of the grammar and vocabulary of language to be elicited. We also found that at 6 months of age, tracking of the speech envelope's phase continues to be universal, taking place for familiar and unfamiliar languages alike. This is in line with what has been observed in adults exhibiting phase-locking to unintelligible speech (Howard & Poeppel, 2010; Pefkou et al., 2017), as well as to non-native language (Song & Iverson, 2018). However, important changes were observed for the tracking of the speech envelope's amplitude at 6 months. At this age, amplitude-encoding of the rhythmically familiar languages ceases to take place (French and Spanish), while it is seemingly enhanced for the rhythmically unfamiliar language (English). A possible explanation for these envelope-tracking differences lies on the previously mentioned theta suppression. We suggest that the rhythm of the incoming speech is processed in the first 500 ms after sentence onset. Once the rhythm is identified, theta oscillations enter a low energy state of entrainment to the rhythmic information (exhibited as suppression), during which phase-encoding (in all languages) and rhythmic discrimination take place. We hypothesize that the theta entrainment triggers the higher frequencies to process other linguistic features (by means of a nesting relationship). Once triggered, higher frequency oscillations preferentially process phonemic units in the native language (and in rhythmically similar ones). This neural commitment of the higher frequencies to the processing of phonemic features reduces their commitment to the amplitude-encoding of the speech envelope, reflected in the absence of amplitude-encoding for the native language (French) and the rhythmically similar unfamiliar language (Spanish). However, for the language rhythmically different from the native language (English), the amplitude of the higher frequencies is not modulated to encode phonemic features, and these frequencies are therefore free to encode the amplitude of speech envelope. This may explain why we observed amplitude envelope tracking to take place only for English at 6 months, since the frequency band envelope tracking is computed over includes frequencies up to 40Hz, i.e. lower gamma. This commitment of the higher frequencies to the processing of phonemic information of their native language is in line with the beginning of the attunement to the native language. It is possible that the processing of the different speech units becomes more efficient with experience, and neural oscillations develop to handle various processes in parallel, which is

why amplitude speech envelope tracking to the native language reappears later in childhood (shown in 9 to 13 year-olds, Abrams et al., 2008).

We also proposed that the observed differences in the amplitude-encoding of the speech envelope could be due to the acoustic shape of the tested English sentences exhibiting more peaks than the French and Spanish sentences due to the high number of stop consonants in them, which would facilitate the correlation to their speech envelopes. This hypothesis would need to be tested by controlling for the acoustic features of the tested stimuli.

5.2.5 Linkage Between Neural Oscillations and Language Outcome

When investigating whether neural oscillations could be predictors of later language outcome, we found that already at birth there was a measure exhibiting this predictive potential. Even though this is far from being clinically relevant, it is scientifically relevant to know that perceptual abilities detectable at birth could tell us something about the trajectory of language acquisition.

The fact that theta activity at birth predicts later language development brings up the question of whether this predictive power in theta develops from prenatal experience with the maternal language, as theta was found to aid in the rhythmic discrimination of the familiar language. If this were the case that would be an indication that even earlier factors (prenatal experience) are influencing language acquisition and future language outcome.

We also found that gamma activity at 6 months was marginally linked to language comprehension at 12 months. We have shown here that gamma activity at 6 months is related to the processing of the native language. This possible link might then indicate that the degree of attunement to the native language exhibited at 6 months, could tell as something about language development.

5.3 LIMITATIONS

Here we have provided new evidence on the neural mechanisms underlying speech processing, language discrimination, and speech envelope tracking. To our knowledge, our study is the first to address this issue in newborns and explore the question systematically in somewhat older infants comparing the native language and rhythmically different unfamiliar languages. And additional strength of our work is our study design in which we tested various conditions (rest, familiar and unfamiliar languages) and different populations (newborns and 6-month-olds). Our large dataset allowed us to use different approaches to analyse it.

Our work however also has some clear limitations. The first limitation pertains to testing time constraints. When designing infant experiments testing sessions need to be kept short. At birth we were able to have recording sessions of approximately half an hour only because the newborns were asleep, otherwise they would had to be much shorter, which was the case for the experiment at 6 months. This time constraints did not allow us to include a resting state condition at 6 months, which therefore will have to be tested in the future, in a separate experiment.

A second challenge (which is also a strength) is the high dimensionality of the dataset, which allows for many types of analyses, but also requires special attention when correcting for multiple comparisons. Infant studies often have low signal to noise ratio datasets, and the use of parametric tests like Bonferroni may hinder us from detecting a small effect. This is the reason why the time-frequency results have not been corrected for multiple comparisons, and we will implement permutation tests (non-parametric) to take into account the multiple comparison problem.

We faced a third limitation when following participants longitudinally. After the newborn group was completed, we had to change the testing location due to administrative reasons. This impacted the return rate of the participants for the recording session at 6 months because some families were less interested in visiting us on the new location, which in most cases was further from their homes.

5.4 FUTURE DIRECTIONS

Here we have shed some light on the developmental origin of neural oscillations and their relationship to speech processing. However, many questions were not addressed here, and some other new questions have emerged.

In the current study we did not investigate the nesting relationship between neural oscillations. This will have to be investigated in future studies to shed light on the hierarchical organization of neural oscillations during infancy. For this purpose, we will compute the cross-frequency phase-amplitude coupling as done by Schroeder & Lakatos, 2009.

In the future we plan to test the resting state condition in 6 month olds to better understand how neural oscillations are modulated during speech processing. Additionally, this resting state condition will contribute in the characterization of the hierarchical organization of spontaneous neural oscillations.

The current work investigated three language conditions were one language was familiar, and a second language was rhythmically familiar. Therefore some of our conclusions cannot be completely disentangled from the familiarity effect. Future studies will require testing similar rhythmic contrasts removing the familiarity effect to be able to draw more general conclusions.

Here we have proposed that speech envelope tracking at birth represents a basic auditory ability independent of language familiarity. New no-language conditions (e.g. natural sounds, animal calls, and music) should be tested to explore how general envelope tracking and neural oscillations are at birth.

We have also argued that the amplitude-encoding of the speech envelope exhibits a U-shape trajectory during development: it is present at birth, not present around 6 and 7 months of age, and then reappears later in childhood (shown in 9 to 13 year-olds, Abrams et al., 2008). This needs to be further investigated testing infants at various age points during the first years of life to confirm and characterized this trajectory.

As it has been previously mentioned, great part of the studies investigating phase-locking to the speech envelope have done so in an indirect way by assessing the inter-trial phase-coherence of their neural responses. To further validate our findings, we plan to also implement this analysis on our dataset.

5.5 GENERAL CONCLUSION

This dissertation contributed to the understanding of the neural mechanisms underlying speech processing and more specifically language discrimination and speech envelope tracking at birth and at 6 months. We showed that neural oscillations already relate to speech processing and envelope tracking at birth, and that they undergo important developmental changes during the first 6 months of life. We have thus highlighted some of the neural mechanisms that may underlie infants' sophisticated speech perception abilities from birth and may lay the foundations for language learning.

REFERENCES

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

Abrams, D. A., Nicol, T., Zecker, S., & Kraus, N. (2008). Right-hemisphere auditory cortex is dominant for coding syllable patterns in speech. *Journal of Neuroscience*, *28*(15), 3958-3965.

Ahissar, E., Nagarajan, S., Ahissar, M., Protopapas, A., Mahncke, H., & Merzenich, M. M. (2001). Speech comprehension is correlated with temporal response patterns recorded from auditory cortex. *Proceedings of the National Academy of Sciences*, *98*(23), 13367-13372.

Allen, E. A., Erhardt, E. B., Damaraju, E., Gruner, W., Segall, J. M., Silva, R. F., ... & Michael, A. M. (2011). A baseline for the multivariate comparison of resting-state networks. *Frontiers in systems neuroscience*, *5*, 2.

Arnal, L. H., Poeppel, D., & Giraud, A. L. (2015). Temporal coding in the auditory cortex. In *Handbook of clinical neurology* (Vol. 129, pp. 85-98). Elsevier.

Astésano, C. (2001). *Rythme et accentuation en français: invariance et variabilité stylistique*. Editions L'Harmattan.

Badin, A. S., Fermani, F., & Greenfield, S. A. (2017). The features and functions of neuronal assemblies: possible dependency on mechanisms beyond synaptic transmission. *Frontiers in neural circuits*, *10*, 114.

Bastiaansen, M., & Hagoort, P. (2006). Oscillatory neuronal dynamics during language comprehension. *Progress in brain research*, *159*, 179-196.

Bastiaansen, M. C., Oostenveld, R., Jensen, O., & Hagoort, P. (2008). I see what you mean: theta power increases are involved in the retrieval of lexical semantic information. *Brain and language*, *106*(1), 15-28.

Bergelson, E., & Swingley, D. (2012). At 6–9 months, human infants know the meanings of many common nouns. *Proceedings of the National Academy of Sciences*, *109*(9), 3253-3258.

Best, C. T., Goldstein, L. M., Nam, H., & Tyler, M. D. (2016). Articulating what infants attune to in native speech. *Ecological Psychology*, *28*(4), 216-261.

Best, C. C., & McRoberts, G. W. (2003). Infant perception of non-native consonant contrasts that adults assimilate in different ways. *Language and speech*, *46*(2-3), 183-216.

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

Bourguignon, M., De Tiege, X., de Beeck, M. O., Ligot, N., Paquier, P., Van Bogaert, P., ... & Jousmäki, V. (2013). The pace of prosodic phrasing couples the listener's cortex to the reader's voice. *Human brain mapping*, *34*(2), 314-326.

Britton, J. W., Frey, L. C., Hopp, J. L., Korb, P., Koubeissi, M. Z., Lievens, W. E., ... & St, E. L. (2016). *Electroencephalography (EEG): An introductory text and atlas of normal and abnormal findings in adults, children, and infants*. American Epilepsy Society, Chicago.

Buzsaki, G. (2006). *Rhythms of the Brain*. Oxford University Press.

Byers-Heinlein, K., Burns, T. C., & Werker, J. F. (2010). The roots of bilingualism in newborns. *Psychological science*, *21*(3), 343-348.

Cheour, M., Ceponiene, R., Lehtokoski, A., Luuk, A., Allik, J., Alho, K., & Näätänen, R. (1998). Development of language-specific phoneme representations in the infant brain. *Nature neuroscience*, *1*(5), 351-353.

Christophe, A., Dupoux, E., Bertoncini, J., & Mehler, J. (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), 1570-1580.

Cristia, A., & Seidl, A. (2011). Sensitivity to prosody at 6 months predicts vocabulary at 24 months. In *BUCLD 35: Proceedings of the 35th annual Boston University Conference on Language Development* (pp. 145-156). Cascadilla Press.

Cristia, A., Seidl, A., Junge, C., Soderstrom, M., & Hagoort, P. (2014). Predicting individual variation in language from infant speech perception measures. *Child development*, *85*(4), 1330-1345.

Cutler, A. (1994). Segmentation problems, rhythmic solutions. *Lingua*, *92*, 81-104.

Cutler, A. (2002). Native listeners. *European Review*, *10*(1), 27-41.

Damoiseaux, J. S., Rombouts, S. A. R. B., Barkhof, F., Scheltens, P., Stam, C. J., Smith, S. M., & Beckmann, C. F. (2006). Consistent resting-state networks across healthy subjects. *Proceedings of the national academy of sciences*, *103*(37), 13848-13853.

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

DeCasper, A. J., & Spence, M. J. (1986). Prenatal maternal speech influences newborns' perception of speech sounds. *Infant Behavior & Development*, *9*(2), 133-150.

Deco, G., Jirsa, V. K., & McIntosh, A. R. (2011). Emerging concepts for the dynamical organization of resting-state activity in the brain. *Nature Reviews Neuroscience*, *12*(1), 43.

Dehaene-Lambertz, G., & Baillet, S. (1998). A phonological representation in the infant brain. *Neuroreport*, *9*(8), 1885-1888.

Dellwo, V. (2006). Rhythm and speech rate: A variation coefficient for deltaC. *Language and Language-Processing: Proceedings of the 38th Linguistic Colloquium*, Frankfurt, Germany, pp. 231–241.

de Haan, M. (2007). Studies in developmental psychology. Infant EEG and event-related potentials. New York, NY, US: Psychology Press.

Denham, K., & Lobeck, A. (2012). *Linguistics for everyone: An introduction*. Cengage Learning.

Ding, N., Melloni, L., Zhang, H., Tian, X., & Poeppel, D. (2016). Cortical tracking of hierarchical linguistic structures in connected speech. *Nature neuroscience*, *19*(1), 158-164.

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

Drullman, R., Festen, J. M., & Plomp, R. (1994a). Effect of temporal envelope smearing on speech reception. *The Journal of the Acoustical Society of America*, *95*(2), 1053-1064.

Drullman, R., Festen, J. M., & Plomp, R. (1994b). Effect of reducing slow temporal modulations on speech reception. *The Journal of the Acoustical Society of America*, *95*(5), 2670-2680.

Eimas, P. D., Siqueland, E. R., Jusczyk, P., & Vigorito, J. (1971). Speech perception in infants. *Science*, *171*(3968), 303-306.

Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D. J., Pethick, S. J. (1994). Variability in early communicative development. *Monographs of the society for research in child development*, *59*(5), 1-173.

Floccia, C., Sambrook, T. D., Delle Luche, C., Kwok, R., Goslin, J., White, L., ... & Mills, D. (2018). I: INTRODUCTION. *Monographs of the Society for Research in Child Development*, *83*(1), 7-29.

Franco, A. R., Pritchard, A., Calhoun, V. D., & Mayer, A. R. (2009). Interrater and intermethod reliability of default mode network selection. *Human brain mapping*, *30*(7), 2293-2303.

Frazier, L., Carlson, K., & Clifton Jr, C. (2006). Prosodic phrasing is central to language comprehension. *Trends in cognitive sciences*, *10*(6), 244-249.

Gerhardt, K. J., Abrams, R. M., & Oliver, C. C. (1990). Sound environment of the fetal sheep. *American journal of obstetrics and gynecology*, *162*(1), 282-287.

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

Gervain, J., Macagno, F., Cogoi, S., Peña, M., & Mehler, J. (2008). The neonate brain detects speech structure. *Proceedings of the National Academy of Sciences*, *105*(37), 14222-14227.

Gervain, J. & Mehler, J. Speech Perception and Language Acquisition in the First Year of Life. *Annual Review of Psychology* **61**, 191-218 (2010).

Gervain, J., & Werker, J. F. (2008). How infant speech perception contributes to language acquisition. *Language and Linguistics Compass*, *2*(6), 1149-1170.

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

Giraud, A. L., Kleinschmidt, A., Poeppel, D., Lund, T. E., Frackowiak, R. S., & Laufs, H. (2007). Endogenous cortical rhythms determine cerebral specialization for speech perception and production. *Neuron*, *56*(6), 1127-1134.

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

Golumbic, E. M. Z., Ding, N., Bickel, S., Lakatos, P., Schevon, C. A., McKhann, G. M., ... & Poeppel, D. (2013). Mechanisms underlying selective neuronal tracking of attended speech at a "cocktail party". *Neuron*, *77*(5), 980-991.

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

Granier-Deferre, C., Lecanuet, J. P., Cohen, H., & Busnel, M. C. (1985). Feasibility of prenatal hearing test. *Acta Oto-Laryngologica*, *99*(sup421), 93-101.

Grieser, D. L., & Kuhl, P. K. (1988). Maternal speech to infants in a tonal language: Support for universal prosodic features in motherese. *Developmental psychology*, *24*(1), 14.

Gross, J., Hoogenboom, N., Thut, G., Schyns, P., Panzeri, S., Belin, P., & Garrod, S. (2013). Speech rhythms and multiplexed oscillatory sensory coding in the human brain. *PLoS biology*, *11*(12), e1001752.

Haarmann, H. J., & Cameron, K. A. (2005). Active maintenance of sentence meaning in working memory: Evidence from EEG coherences. *International Journal of Psychophysiology*, *57*(2), 115-128.

He, B. J. (2014). Scale-free brain activity: past, present, and future. *Trends in cognitive sciences*, *18*(9), 480-487.

Hoareau, M. (2019). The influence of parental input and early production skills on speech perception abilities and vocabulary acquisition during the first year of life (doctoral dissertation). Université Paris Descartes, Paris, France.

Homae, F., Watanabe, H., Nakano, T., & Taga, G. (2011). Large-scale brain networks underlying language acquisition in early infancy. *Frontiers in psychology*, *2*, 93.

Howard, M. F., & Poeppel, D. (2010). Discrimination of speech stimuli based on neuronal response phase patterns depends on acoustics but not comprehension. *Journal of neurophysiology*, *104*(5), 2500-2511.

Howard, M. F., & Poeppel, D. (2012). The neuromagnetic response to spoken sentences: co-modulation of theta band amplitude and phase. *Neuroimage*, *60*(4), 2118-2127.

Hurtado, J. M., Rubchinsky, L. L., & Sigvardt, K. A. (2004). Statistical method for detection of phase-locking episodes in neural oscillations. *Journal of neurophysiology*, *91*(4), 1883-1898.

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

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

Jusczyk, P. W., Hohne, E. A., & Bauman, A. (1999). Infants' sensitivity to allophonic cues for word segmentation. *Perception & psychophysics*, *61*(8), 1465-1476.

Kalashnikova, M., Peter, V., Di Liberto, G. M., Lalor, E. C., & Burnham, D. (2018). Infant-directed speech facilitates seven-month-old infants' cortical tracking of speech. *Scientific reports*, 8(1), 13745

Kern, S. (2007). Lexicon development in French-speaking infants. *First Language*, *27*(3), 227-250.

Kubanek, J., Brunner, P., Gunduz, A., Poeppel, D., & Schalk, G. (2013). The tracking of speech envelope in the human cortex. *PloS one*, *8*(1), e53398.

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

Kuhl, P. K., Conboy, B. T., Coffey-Corina, S., Padden, D., Rivera-Gaxiola, M., & Nelson, T. (2007). Phonetic learning as a pathway to language: new data and native language magnet theory expanded (NLM-e). *Philosophical Transactions of the Royal Society B: Biological Sciences*, *363*(1493), 979-1000.

Kuhl, P. K., Conboy, B. T., Padden, D., Nelson, T., & Pruitt, J. (2005). Early speech perception and later language development: Implications for the" critical period". *Language learning and development*, *1*(3-4), 237-264.

Ladefoged, P., & Johnson, K. (2011). *A course in phonetics*. Cengage Learning.

Lakatos, P., Shah, A. S., Knuth, K. H., Ulbert, I., Karmos, G., & Schroeder, C. E. (2005). An oscillatory hierarchy controlling neuronal excitability and stimulus processing in the auditory cortex. *Journal of neurophysiology*, *94*(3), 1904-1911.

Lakatos, P., Szilágyi, N., Pincze, Z., Rajkai, C., Ulbert, I., & Karmos, G. (2004). Attention and arousal related modulation of spontaneous gamma-activity in the auditory cortex of the cat. *Cognitive brain research*, *19*(1), 1-9.

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

Lehongre, K., Ramus, F., Villiermet, N., Schwartz, D., & Giraud, A. L. (2011). Altered low-gamma sampling in auditory cortex accounts for the three main facets of dyslexia. *Neuron*, *72*(6), 1080-1090.

Lewis, A. G., & Bastiaansen, M. (2015). A predictive coding framework for rapid neural dynamics during sentence-level language comprehension. *Cortex*, *68*, 155-168.

Luo, H., & Poeppel, D. (2007). Phase patterns of neuronal responses reliably discriminate speech in human auditory cortex. *Neuron*, *54*(6), 1001-1010.

Magne, C., Aramaki, M., Astesano, C., Gordon, R. L., Ystad, S., Farner, S., ... & Besson, M. (2004). Comparison of rhythmic processing in language and music: An interdisciplinary approach. *JMM: The Journal of Music and Meaning*, *3*, 2004-2005.

Mahmoudzadeh, M., Dehaene-Lambertz, G., Fournier, M., Kongolo, G., Goudjil, S., Dubois, J., ... & Wallois, F. (2013). Syllabic discrimination in premature human infants prior to complete formation of cortical layers. *Proceedings of the National Academy of Sciences*, *110*(12), 4846-4851.

Mai, G., Minett, J. W., & Wang, W. S. Y. (2016). Delta, theta, beta, and gamma brain oscillations index levels of auditory sentence processing. *Neuroimage*, *133*, 516-528.

Marina Nespor & Irene Vogel (1986). Prosodic phonology. Dordrecht: Foris Publications. Pp. xiv 327. *Phonology*, 5(1), 161-168.

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

Maye, J., Werker, J. F., & Gerken, L. (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition*, *82*(3), B101-B111.

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

Meyer, L. (2018). The neural oscillations of speech processing and language comprehension: state of the art and emerging mechanisms. *European Journal of Neuroscience*, *48*(7), 2609-2621.

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

Molnar, M., Gervain, J., & Carreiras, M. (2013). Within-rhythm class native language discrimination abilities of Basque-Spanish monolingual and bilingual infants at 3.5 months of age. *Infancy*, *19*(3), 326-337.

Moon, C., Cooper, R. P., & Fifer, W. P. (1993). Two-day-olds prefer their native language. *Infant behavior and development*, *16*(4), 495-500.

Moon, C., Lagercrantz, H., & Kuhl, P. K. (2013). Language experienced in utero affects vowel perception after birth: A two-country study. *Acta Paediatrica*, *102*(2), 156-160.

Moon, I. J., & Hong, S. H. (2014). What is temporal fine structure and why is it important?. *Korean journal of audiology*, *18*(1), 1-7.

Morillon, B., Lehongre, K., Frackowiak, R. S., Ducorps, A., Kleinschmidt, A., Poeppel, D., & Giraud, A. L. (2010). Neurophysiological origin of human brain asymmetry for speech and language. *Proceedings of the National Academy of Sciences*, *107*(43), 18688-18693

Morillon, B., Liégeois-Chauvel, C., Arnal, L. H., Bénar, C. G., & Giraud, A. L. (2012). Asymmetric function of theta and gamma activity in syllable processing: an intra-cortical study. *Frontiers in psychology*, *3*, 248.

Nazzi, T., Bertoncini, J., & Mehler, J. (1998). Language discrimination by newborns: toward an understanding of the role of rhythm. *Journal of Experimental Psychology: Human perception and performance*, *24*(3), 756-766.

Nazzi, T., Jusczyk, P. W., & Johnson, E. K. (2000). Language discrimination by English-learning 5-month-olds: Effects of rhythm and familiarity. *Journal of Memory and Language*, *43*(1), 1-19.

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

Nourski, K. V., Reale, R. A., Oya, H., Kawasaki, H., Kovach, C. K., Chen, H., ... & Brugge, J. F. (2009). Temporal envelope of time-compressed speech represented in the human auditory cortex. *Journal of Neuroscience*, *29*(49), 15564-15574.

Ortiz-Barajas, M. C., & Gervain, J. (submitted). The role of prenatal experience and basic auditory mechanisms in the development of language. *Minnesota Symposia on Child Psychology*, Volume 40: Human Communication: Origins, Mechanisms and Functions, Editors: M. Sera and M. Koenig. Wiley E-book, Hoboken, NJ.

Parise, E., & Csibra, G. (2013). Neural responses to multimodal ostensive signals in 5-month-old infants. *PloS one*, *8*(8), e72360.

Peelle, J. E., Gross, J., & Davis, M. H. (2013). Phase-locked responses to speech in human auditory cortex are enhanced during comprehension. *Cerebral cortex*, *23*(6), 1378-1387.

Pefkou, M., Arnal, L. H., Fontolan, L., & Giraud, A. L. (2017). θ-Band and β-band neural activity reflects independent syllable tracking and comprehension of time-compressed speech. *Journal of Neuroscience*, *37*(33), 7930-7938.

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

Peña, M., & Melloni, L. (2012). Brain oscillations during spoken sentence processing. *Journal of cognitive neuroscience*, *24*(5), 1149-1164.

Peña, M., Pittaluga, E., & Mehler, J. (2010). Language acquisition in premature and full-term infants. *Proceedings of the National Academy of Sciences*, *107*(8), 3823-3828.

Pujol, R., & Lavigne-Rebillard, M. (1991). Development and plasticity of the human auditory system. *Acta oto-laryngologica*, pp 13-22.

Querlue, D., & Renard, K., (1981). Les perceptions auditives du foetus humain. Medicine & Hygiene, 39, 2101-2110

Querlue, D., Renard, K., & Crepin, G. (1981). Perception auditive et reactivite foetale aux stimulations sonores. Journal de Gynecologie Obstetrique et Biologie de la Reproduction, 10, 307-314

Ramus, F., & Ahissar, M. (2012). Developmental dyslexia: The difficulties of interpreting poor performance, and the importance of normal performance. *Cognitive neuropsychology*, *29*(1-2), 104-122.

Ramus, F., Hauser, M. D., Miller, C., Morris, D., & Mehler, J. (2000). Language discrimination by human newborns and by cotton-top tamarin monkeys. *Science*, *288*(5464), 349-351.

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

Rivera-Gaxiola, M., Klarman, L., Garcia-Sierra, A., & Kuhl, P. K. (2005a). Neural patterns to speech and vocabulary growth in American infants. *NeuroReport*, *16*(5), 495-498.

Rivera-Gaxiola, M., Silva-Pereyra, J., & Kuhl, P. K. (2005b). Brain potentials to native and non-native speech contrasts in 7-and 11-month-old American infants. *Developmental science*, *8*(2), 162-172.

Rosen, S. (1992). Temporal information in speech: acoustic, auditory and linguistic aspects. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *336*(1278), 367-373.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, *274*(5294), 1926-1928.

Saffran, J. R., & Thiessen, E. D. (2003). Pattern induction by infant language learners. *Developmental psychology*, *39*(3), 484-494.

Sansavini, A., Bertoncini, J., & Giovanelli, G. (1997). Newborns discriminate the rhythm of multisyllabic stressed words. *Developmental Psychology*, *33*(1), 3-11.

Sato, H., Hirabayashi, Y., Tsubokura, H., Kanai, M., Ashida, T., Konishi, I., ... & Maki, A. (2012). Cerebral hemodynamics in newborn infants exposed to speech sounds: A whole-head optical topography study. *Human brain mapping*, *33*(9), 2092-2103.

Schroeder, C. E., & Lakatos, P. (2009). Low-frequency neuronal oscillations as instruments of sensory selection. *Trends in neurosciences*, *32*(1), 9-18.

Shannon, R. V., Zeng, F. G., Kamath, V., Wygonski, J., & Ekelid, M. (1995). Speech recognition with primarily temporal cues. *Science*, *270*(5234), 303-304.

Shaywitz, S. E. (1998). Dyslexia. *New England Journal of Medicine*, *338*(5), 307-312.

Snyder, A. Z., & Raichle, M. E. (2012). A brief history of the resting state: the Washington University perspective. *Neuroimage*, *62*(2), 902-910.

Song, J., & Iverson, P. (2018). Listening effort during speech perception enhances auditory and lexical processing for non-native listeners and accents. *Cognition*, *179*, 163-170.

Stefanics, G., Háden, G. P., Sziller, I., Balázs, L., Beke, A., & Winkler, I. (2009). Newborn infants process pitch intervals. *Clinical Neurophysiology*, *120*(2), 304-308.

Steinhauer, K., Connolly, J. F. (2008). Event-related potentials in the study of language. *Concise encyclopedia of brain and language*, 191-203.

Steriade, M., Contreras, D., Amzica, F., & Timofeev, I. (1996). Synchronization of fast (30-40 Hz) spontaneous oscillations in intrathalamic and thalamocortical networks. *Journal of Neuroscience*, *16*(8), 2788-2808.

Sun, J., Hong, X., & Tong, S. (2011). A comparison of surrogate tests for phase synchronization analysis of neural signals. In *Proc. Asia-Pacific Signal Inf. Process. Assoc. Annu. Summit Conf* (Vol. 203).

Tincoff, R., & Jusczyk, P. W. (1999). Some beginnings of word comprehension in 6-month-olds. *Psychological Science*, *10*(2), 172-175.

Tincoff, R., & Jusczyk, P. W. (2012). Six-month-olds comprehend words that refer to parts of the body. *Infancy*, *17*(4), 432-444.

Tomblin, J. B., Records, N. L., Buckwalter, P., Zhang, X., Smith, E., & O'Brien, M. (1997). Prevalence of specific language impairment in kindergarten children. *Journal of speech, language, and hearing research*, *40*(6), 1245-1260.

Tóth, B., Urbán, G., Háden, G. P., Márk, M., Török, M., Stam, C. J., & Winkler, I. (2017). Large-scale network organization of EEG functional connectivity in newborn infants. *Human brain mapping*, *38*(8), 4019-4033.

Tsao, F. M., Liu, H. M., & Kuhl, P. K. (2004). Speech perception in infancy predicts language development in the second year of life: A longitudinal study. *Child development*, *75*(4), 1067-1084.

Vander Ghinst, M., Bourguignon, M., de Beeck, M. O., Wens, V., Marty, B., Hassid, S., ... & Goldman, S. (2016). Left superior temporal gyrus is coupled to attended speech in a cocktail-party auditory scene. *Journal of Neuroscience*, *36*(5), 1596-1606.

Varnet, L., Ortiz-Barajas, M. C., Erra, R. G., Gervain, J., & Lorenzi, C. (2017). A cross-linguistic study of speech modulation spectra. *The Journal of the Acoustical Society of America*, *142*(4), 1976-1989.

Vince, M. A., & Armitage, S. E. (1980). Sound stimulation available to the sheep foetus. *Reproduction Nutrition Développement*, *20*(3B), 801-806.

Vouloumanos, A., & Werker, J. F. (2007). Listening to language at birth: Evidence for a bias for speech in neonates. *Developmental science*, *10*(2), 159-164.

Walker, D., Grimwade, J., & Wood, C. (1971). Intrauterine noise: A component of the fetal environment. *American Journal of Obstetrics and Gynecology*, *109*(1), 91-95.

Wang, X. J. (2010). Neurophysiological and computational principles of cortical rhythms in cognition. *Physiological reviews*, *90*(3), 1195-1268.

Weikum, W. M., Oberlander, T. F., Hensch, T. K., & Werker, J. F. (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), 17221-17227.

Werker, J. F., & Tees, R. C. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant behavior and development*, *7*(1), 49-63.