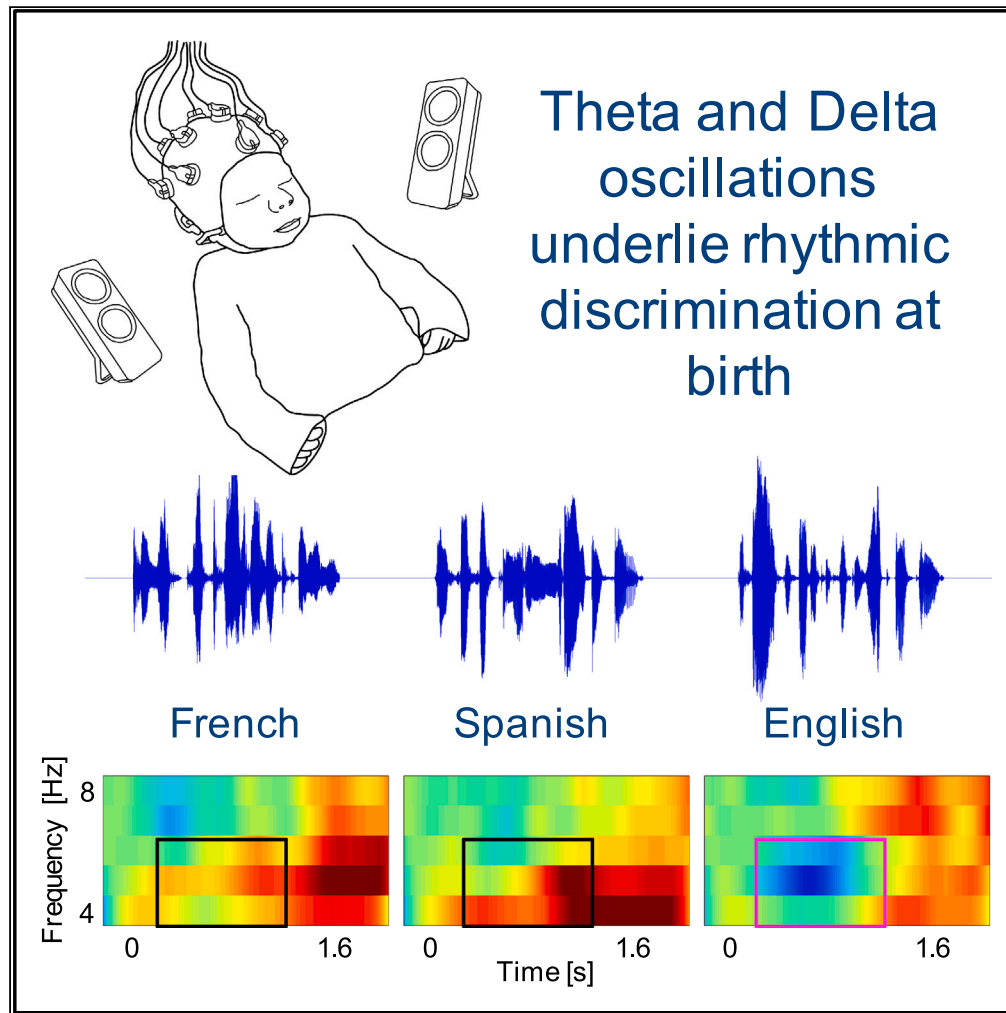


Article

Neural oscillations and speech processing at birth



Maria Clemencia
Ortiz-Barajas,
Ramón Guevara,
Judit Gervain

maria.c.ortizb@gmail.com

Highlights

Neural oscillations support
speech processing from
birth

Delta and theta oscillations
underlie rhythmic
language discrimination at
birth

Theta activity exhibits
stimulation after-effects
following speech
stimulation

Neural oscillations are
hierarchically organized
into a nesting relation at
birth

Ortiz-Barajas et al., iScience 26,
108187
November 17, 2023 © 2023 The
Authors.
[https://doi.org/10.1016/
j.isci.2023.108187](https://doi.org/10.1016/j.isci.2023.108187)

Article

Neural oscillations and speech processing at birth

Maria Clemencia Ortiz-Barajas,^{1,4,*} Ramón Guevara,² and Judit Gervain^{1,3}

SUMMARY

Are neural oscillations biologically endowed building blocks of the neural architecture for speech processing from birth, or do they require experience to emerge? In adults, delta, theta, and low-gamma oscillations support the simultaneous processing of phrasal, syllabic, and phonemic units in the speech signal, respectively. Using electroencephalography to investigate neural oscillations in the newborn brain we reveal that delta and theta oscillations differ for rhythmically different languages, suggesting that these bands underlie newborns' universal ability to discriminate languages on the basis of rhythm. Additionally, higher theta activity during post-stimulus as compared to pre-stimulus rest suggests that stimulation after-effects are present from birth.

INTRODUCTION

How does our brain process speech so rapidly and automatically that we can hold conversations without much effort? The human brain must extract the linguistic units that are present in speech (i.e., syllables, words, phrases) to process them and integrate them into a linguistic representation allowing us to interpret speech. This process occurs very rapidly, as the brain smoothly processes the different units of speech simultaneously.¹ One neural mechanism that has been suggested to be involved in speech parsing and decoding is that of neural oscillations,^{2–6} i.e., the rhythmic brainwaves generated by the synchronous activity of neuronal assemblies, which can be coupled intrinsically, i.e., in the absence of stimulation, or by a common input.⁷

The role of neural oscillations in human language processing has received increasing attention. Electrophysiological recordings from the human auditory cortex show that oscillatory activity in the delta, theta, alpha and low-gamma bands may be observed during resting state.^{3,8,9} During speech processing, the auditory cortex is stimulated, and resting state oscillations are transformed into temporally structured oscillations.³ Thus neural oscillations in the auditory cortex provide the infrastructure to parse and decode continuous speech into its constituent units at different levels synchronously, as it has been argued theoretically³ and shown empirically.^{2,6,10} More specifically, 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 simultaneously into (sub)phonemic (30–50 Hz), syllabic (4–7 Hz) and phrasal (1–2 Hz) units, respectively. This hypothesis is partly based on the close correspondence between the frequencies at which these oscillations operate and the time scales of the corresponding speech units. Importantly, these oscillations are organized into a hierarchy or a nesting/entrainment relation, whereby the lower frequency oscillations enslave the higher frequency ones, as has been shown both in humans (theta–gamma coupling¹¹) and animals (delta–theta coupling¹²). Specifically, the amplitude of higher frequency oscillations is modulated by the phase of lower frequency oscillations.^{12–14} This hierarchical organization explains how we can process speech units of different sizes simultaneously, leading to the formulation of the multi-timescale model of speech perception.³

Neural oscillations are now well established as a key mechanism of speech and language processing in adults. Indeed, in adults, neural oscillations have been shown to synchronize with phonemes,¹⁵ syllables¹⁶ and larger prosodic phrases,^{10,17} and the extent of this synchronization has been shown to predict intelligibility.^{4,10,16,18,19} Furthermore, neural entrainment has been causally related to perception, as transcranially induced oscillations modulate auditory detection thresholds²⁰ and speech intelligibility²¹ at the behavioral level. Oscillations have also been observed to play a role in establishing the syntactic parsing of sentences, as oscillatory activity at the frequency of syntactic phrases has been found to arise only in listeners who understand the linguistic stimuli presented,¹⁰ as well as in semantic processing, with power in the gamma range increasing for semantically coherent and decreasing for incoherent linguistic stimuli.²² In addition to these bottom-up, signal-driven oscillations, the beta band has recently been identified as providing top-down, predictive information derived from the listener's linguistic knowledge.²³

Despite the wealth of evidence in adults, the developmental origins of neural oscillations and their contribution to language development remain largely unexplored. As oscillations are present in the auditory cortex of non-human animals,¹² they may also be operational in humans from birth. It is also possible, however, that oscillations and their coupling emerge through experience with speech and language, or with the auditory environment in general.

¹Integrative Neuroscience and Cognition Center, CNRS & Université Paris Cité, 45 rue des Saints-Pères, 75006 Paris, France

²Department of Physics and Astronomy, University of Padua, Via Marzolo 8, 35131 Padua, Italy

³Department of Developmental and Social Psychology, University of Padua, Via Venezia 8, 35131 Padua, Italy

⁴Lead contact

*Correspondence: mariac.ortizb@gmail.com

<https://doi.org/10.1016/j.isci.2023.108187>



A few existing studies have investigated neural oscillations and their relationship with some aspects of speech processing in older infants. Theta oscillations have been found to be sensitive to syllable frequency (frequent vs. infrequent) at 6 months, and to stimulus familiarity (native vs. non-native syllables) at 12 months, suggesting that changes in theta sensitivity reflect a shift in the focus of attention from frequency at 6 months to categories at 12 months.²⁴ Theta oscillations have also been found to support language discrimination at 4.5 months, reflecting rhythmic discrimination in monolinguals, and within rhythmic class discrimination in bilinguals.²⁵ Moreover, gamma oscillations have been found to be sensitive to visual and auditory modalities of ostensive communication at 5 months,²⁶ as well as sensitive to stimulus familiarity (native vs. non-native syllables) at 6 months (low-gamma)²⁷ and 12 months (high-gamma).²⁸ Gamma oscillations have also been found to underlie the discrimination of rhythmically similar languages at 6 months in full-term and at 9 months in pre-term infants, suggesting that these oscillations are under maturational control.²⁹ Neural oscillations have also been found to synchronize with the amplitude and phase of the speech envelope of familiar and unfamiliar languages at birth,³⁰ while at around 6–7 months amplitude tracking is absent for the native language^{30,31} and present for unfamiliar languages.³⁰ These studies suggest that neural oscillations and their role in speech processing change with brain maturation and language exposure.³²

The human brain also extracts regularities in rhythmic non-speech auditory stimuli. Adults and infants' neural activity entrains to isochronous rhythmic beats,^{33,34} and in adults, also to imagined meters (i.e., at $f/2$ and $f/3$).³³ Beta oscillations are sensitive to beat onset, and reflect a neural mechanism predicting the next stimulus in adults,^{35,36} and children.³⁷ This neural ability to entrain to rhythm could explain why neural oscillations have been found to synchronize with speech, an auditory signal known to contain (quasi-)periodic information.

No study to date has investigated neural oscillations at the beginning of extrauterine life, their developmental origins thus remain critically unexplored. Yet human newborns have sophisticated speech perception abilities. Some of these are universal and broadly based, allowing them to acquire any language. For instance, they can discriminate rhythmically different, but not rhythmically similar languages, even if those are unfamiliar to them.^{38,39} Thus, newborns prenatally exposed to French are able to discriminate Spanish from English, because these two languages have different rhythmic properties.⁴⁰ Rhythmic discrimination is fundamental for language development⁴¹ and underlies multilinguals' ability to identify dual input early on.⁴² Interestingly, neonates also show speech perception abilities shaped by prenatal experience with the language(s) spoken by the mother during the last trimester of pregnancy. Newborns' prenatal experience with speech mainly consists of language prosody, i.e., rhythm and melody, because maternal tissues filter out the higher frequencies necessary for the identification of individual phonemes, but preserve the low-frequency components that carry prosody.⁴³ On the basis of this experience, newborns are able to recognize their native language, and prefer it over other languages,^{41,44} and may even modulate the contours of their cries to match the prosody of their native language.⁴⁵

Do neural oscillations already support these early neonatal abilities or do they emerge as infants gain experience with language over developmental time? To address these critical unanswered questions, here we investigated neural oscillations in human newborns during the speech processing of familiar and unfamiliar languages as well as at rest before and after stimulation. Specifically, we seek to address (i) whether speech stimuli trigger similar neural oscillations at birth as in adults, (ii) whether prenatal experience with the maternal language modulates neural oscillations, (iii) whether neural oscillations are already hierarchically organized, i.e., nested, at birth and (iv) whether the organization of auditory oscillations triggered by language stimuli are maintained even after stimulation. Ours is the first study to investigate the hierarchical organization of neural oscillations, and their response to speech at birth.

We analyzed data from 40 full-term, healthy newborns (mean age: 2.55 days; range: 1–5 days; 17 females), born to French monolingual mothers, within their first 5 days of life. Their experience with speech was, therefore, mostly prenatal. We tested newborns at rest (pre- and post-stimulation) and while listening to repetitions of naturally spoken sentences (~ 2 s in duration) in three languages: their native language, i.e., the language heard prenatally, French, a rhythmically similar unfamiliar language, Spanish, and a rhythmically different unfamiliar language, English (Figure 1C) (see Audios S1–S9). While these languages also differ in other phonological properties such as phoneme repertoire or syllable structure, newborns have not been shown to rely on features other than rhythmic class for language discrimination. We thus focus on the rhythmic differences without denying that other properties may also play a role.

Our predictions are the following. (i) If the prenatally heard low-pass filtered, mainly prosodic speech signal already shapes the neural mechanisms of speech processing, then we expect the low frequency neural oscillations, delta and theta, known to encode prosodic information in adults, and not the higher frequency gamma band, to be present and enhanced at birth during the language conditions. (ii) Similarly, as rhythm is carried by prosody, which is encoded by lower frequency oscillations, the neural signature of the behaviorally attested ability to discriminate rhythmically different languages, if found, is expected to occur in the lower frequency ranges, delta and theta. (iii) For the within-rhythm class discrimination (French vs. Spanish) we do not expect neural oscillations to differ, because behavioral studies have shown that they are not discriminated at birth.³⁸ Therefore, the within-rhythmic class comparison is exploratory and leverages the fact that neural imaging may be more sensitive than behavioral measures. (iv) Additionally, we predict that post-stimulus resting state may be modified by stimulation, as has been found for functional connectivity after language stimulation in 3-month-olds.⁴⁶

RESULTS

Time-frequency responses

We filtered the continuous EEG data, segmented it into epochs time-locked to the beginning of each sentence, and rejected epochs containing artifacts (see STAR methods). We then subjected each non-rejected epoch to time-frequency analysis to uncover oscillatory responses using the MATLAB toolbox 'WTools'.²⁶ This toolbox performs a continuous wavelet transform of each epoch using Morlet wavelets (number of cycles was 3.5). The time-frequency responses were averaged across non-rejected epochs for each condition and channel separately.

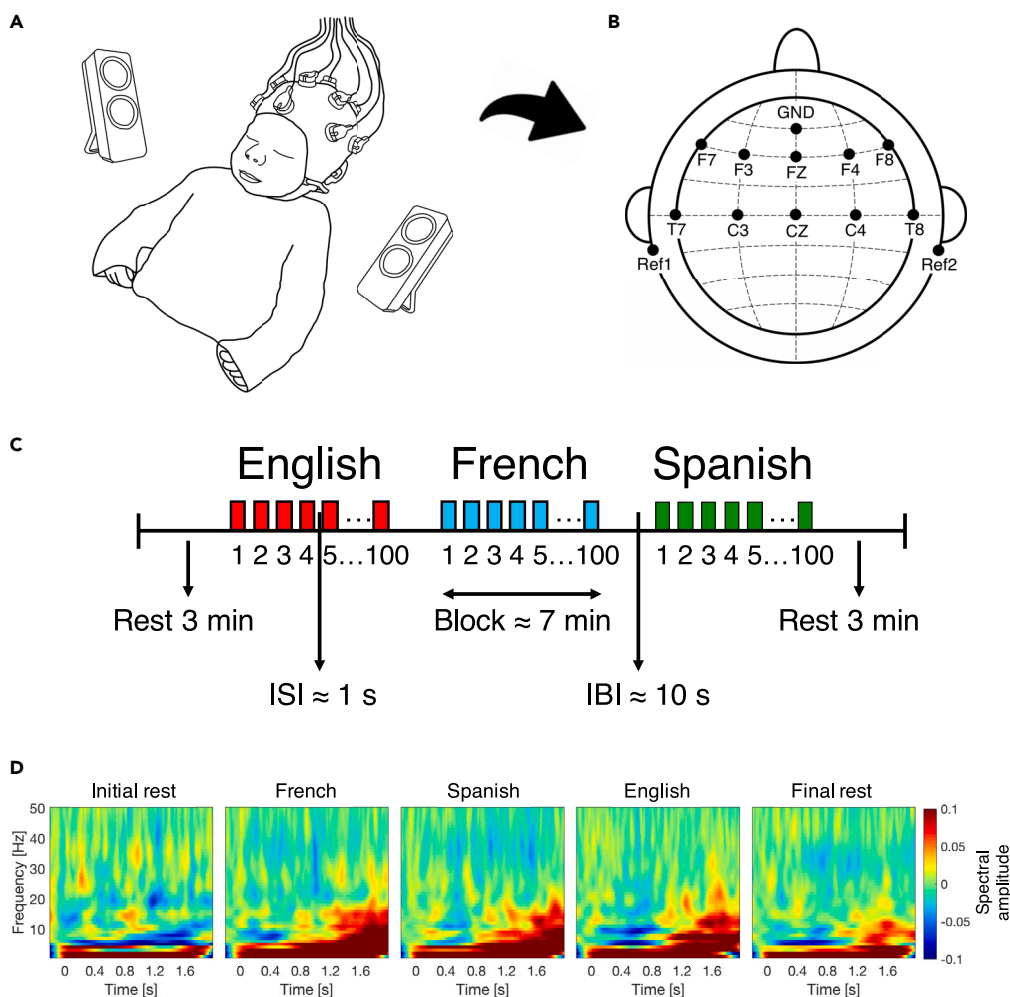


Figure 1. EEG experimental design and time-frequency responses

(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. Figure adapted from Ortiz Barajas et al., 2021.³⁰

(D) The average time-frequency response for each tested condition at channel F8. The time-frequency maps illustrate the mean spectral amplitude per condition from 1 to 50 Hz. The color bar to the right of the figure shows the spectral amplitude scale for all maps. See also Figure S1.

Figure 1D displays the group mean time-frequency response for each condition at channel F8 as an example (Figure S1 displays the response at other channels). We submitted the time-frequency responses to various statistical analyses to compare newborns' oscillatory activation (i) across language conditions, (ii) between pre-stimulus rest and speech and (iii) between pre- and post-stimulus rest.

Language differences

We were first interested in determining whether neural oscillations are modulated by language familiarity and/or rhythm. To assess possible differences in activation across languages, we submitted the time-frequency responses to the three language conditions to permutation testing involving repeated measures ANOVAs with the within-subject factor Language (French/Spanish/English). (Figure S2A presents *P*-maps for this comparison). Chance distribution was established and multiple comparisons were corrected for by permuting the language labels in the original dataset (see STAR methods and Figure S2C, which illustrates the chance distributions for the significant clusters). These analyses yielded activation differences across languages in the delta and theta bands, discussed later in discussion. Table 1 (panel A) summarizes the activation differences, and Figure S2B highlights the time-frequency windows where these differences take place.

Delta band. Figure 2 displays the significant differences observed between languages in the delta band (1–3 Hz). Frontal channels F7 [$F(2,78) = 1626.3$, $p = 0.032$,] and F8 [$F(2,78) = 1380.8$, $p = 0.043$] (Figure S2C) exhibit higher activation when processing the native language,

Table 1. Summary of neural activation differences across conditions

Panel A: Differences across languages	
Neural Oscillations	French and Spanish > English
Delta [1–3 Hz]	F7, F8
Theta [4–8 Hz]	F4
Panel B: Differences between the initial resting state and speech	
Neural Oscillations	Speech > Initial rest
Delta [1–3 Hz]	T8
Theta [4–8 Hz]	F7, F4, F8, T7, C4, T8
Panel C: Differences between the initial and final resting state	
Neural Oscillations	Final rest > Initial rest
Theta [4–8 Hz]	T7, T8

French and the rhythmically similar unfamiliar language, Spanish than the rhythmically different unfamiliar language, English at 2 Hz during the second half of the sentences. The maximum effect size, partial eta-squared (η_p^2), for the cluster in channel F7 is 0.1158, and for the cluster in channel F8 is 0.1688.

Theta band. Figure 3 displays the significant differences observed between languages in the theta band (4–8 Hz). Frontal channel F4 [$F(2,78) = 1519.4$, $p = 0.016$] (Figure S2C) exhibits higher activation when processing French and Spanish than English, mainly at 5 Hz during the first half of the sentences. The maximum effect size, partial eta-squared (η_p^2), for the cluster in channel F4 is 0.1448.

The previous results were obtained for a sample size of 40 participants, who listened to sentences of different durations (see STAR methods). To be able to compare the neural responses obtained for different sentence durations, we chopped the EEG epochs to the length of the shortest sentence. To further support our results, we have conducted additional analyses on the subset of participants ($n = 14$) who listened to the longest sentences (Set 3). Results from these analyses are presented in the supplementary information (Figures S3 and S4), and they are congruent with our previous findings. Language differences are mainly found in the theta band, where neural activation is higher for Spanish than for French and English toward the first half of the sentences (channels F8 [$F(2,26) = 1896.9$, $p = 0.04$] and T8 [$F(2,26) = 451.9$, $p = 0.04$]); and higher for French than for Spanish and English toward the end of the sentences (channel T7 [$F(2,26) = 1600.1$, $p = 0.03$]). The maximum effect size, partial eta-squared (η_p^2), for the cluster in channel F8 is 0.3516, in channel T7 is 0.4599, and in channel T8 is 0.2823.

These results establish that already at birth, neural oscillations in the human brain are sensitive to differences across languages. Oscillations successfully distinguish between rhythmically different languages, English, a stress-timed language, from French and Spanish, two syllable-timed languages,⁴⁰ underlying newborns' behaviorally well established rhythmic discrimination ability.^{38,39} These differential responses occur in the slower bands, delta and theta, while faster oscillations are less influenced by language familiarity or rhythm. Moreover, no significant differences in activation were found between the rhythmically similar languages (French and Spanish), which is in line with our prediction based on findings from behavioral studies.

Resting state vs. speech processing

To assess neural oscillations during rest at birth, and contrast it with oscillatory activity during speech processing, we compared the time-frequency responses from the initial resting period to those from the speech stimulation. In order to evaluate speech processing in a general, non-language specific way, we averaged the time-frequency responses from the three languages to obtain what we call hereafter a speech stimulation response. Figure 4 displays the time-frequency responses to resting state and speech in channel T8 as an example. We submitted the time-frequency responses from the initial resting state condition and the speech stimulation condition to permutation tests involving two-tailed paired-samples t-tests (Figure S5A presents P-maps for this comparison). Table 1 (panel B) summarizes the activation differences, and Figure S5B highlights the time-frequency windows where these differences take place. As predicted, in the lower frequency bands (delta and theta) activation is higher during speech processing than during the initial resting state period: at T8 for delta [$t(35) = 1033.6$, $p = 0.016$]; and at F7 [$t(35) = 1703.3$, $p = 0.006$], F4 [$t(35) = 2242.5$, $p = 0.005$], F8 [$t(35) = 3865.8$, $p = 0$], T7 [$t(35) = 1780.1$, $p = 0.001$], C4 [$t(35) = 1624.5$, $p = 0.005$], T8 [$t(35) = 3212.9$, $p = 0.001$] for theta. The maximum effect size, Cohen's d , for the cluster in the delta band (channel T8) is 0.7239; while for the clusters in the theta band it is 0.8951 at F7, 0.8796 at F4, 1.2894 at F8, 1.1028 at T7, 0.9564 at C4, and 0.9127 at T8. We also compared the time-frequency responses from the initial resting period to those from each language separately (Figure S6). Results for the individual languages are similar to those found for the speech-general response.

Stimulation after-effects

The newborn brain is highly plastic and language experience shapes speech perception abilities and their neural correlates already very early on in development, i.e., prenatally and over the first months of life, attuning the infant's perceptual system and the brain's language network

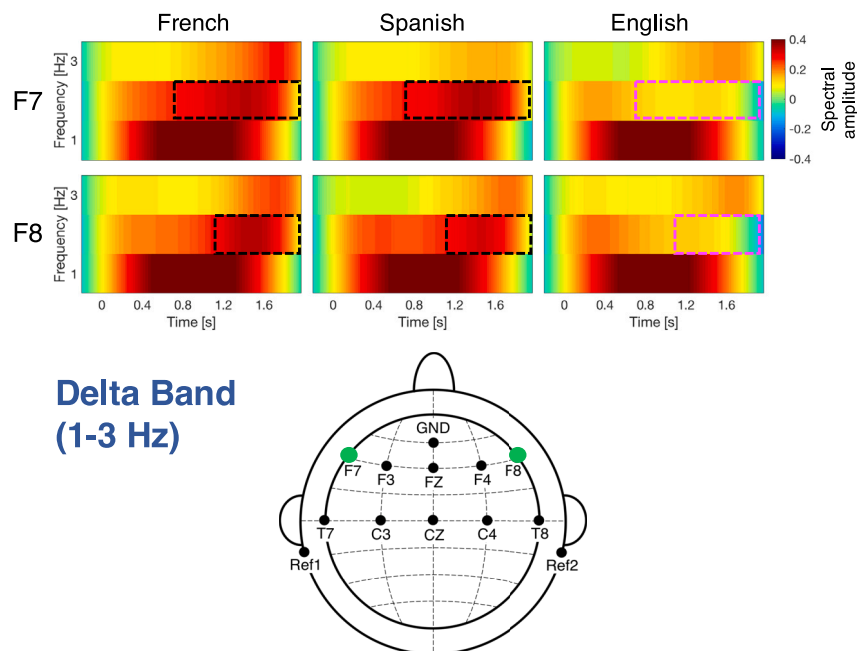


Figure 2. Time-frequency response in the delta band (1–3 Hz) for channels exhibiting significant differences across languages

Channels F7 and F8 have significantly higher activation for French and Spanish than for English in the time and frequency ranges indicated by the black rectangular boxes (magenta boxes indicate significantly lower activation). See also Figure S2.

to the native language.^{47,48} Exploring newborns' oscillatory responses after speech stimulation can reveal the neural mechanisms underlying learning and perceptual attunement to the native language. To assess whether stimulation after-effects are already present at birth, we submitted the time-frequency responses from the initial and final resting state periods to permutation tests involving two-tailed paired-samples t-tests (Figure S7A presents *P*-maps for this comparison). Table 1 (panel C) summarizes the activation differences between the two resting state conditions, and Figure S7B highlights the time-frequency windows where these differences take place. Figure 4 illustrates these differences in channel T8 as an example. In the theta band, oscillatory power is higher during the final resting state period than during the initial one in channels T7 [$t(35) = 1692.9$, $p = 0.01$] and T8 [$t(35) = 820.9$, $p = 0.01$]. The maximum effect size, Cohen's *d*, for the cluster in channel T7 is 0.8804, and for the cluster in channel T8 is 0.7743.

When comparing the initial and final resting periods (Table 1, panel C), differences are similar to those observed between initial rest and speech (Table 1, panel B). Neural activity after a long period (21 min) of speech stimulation, therefore, does not immediately return to baseline.

Nesting

To investigate whether neural oscillations are already hierarchically organized at birth, which is necessary for processing linguistic units at different levels simultaneously, we evaluated the coupling between the phase of low-frequency oscillations, and the amplitude of high-frequency oscillations in our newborn data. To do so, we subjected each non-rejected epoch to a continuous wavelet transform using the MATLAB toolbox 'WTools'²⁶ to obtain the phase and amplitude of the oscillations for each trial (see STAR methods). We then computed the modulation index⁴⁹: we first obtained the modulation index as a function of analytic phase (1–30 Hz) and analytic amplitude (1–50 Hz) for each non-rejected epoch, channel and condition separately, and then we averaged it across non-rejected epochs. Figure 5 displays the modulation index for each condition in channel F8 as an example. Larger modulation index values indicate stronger cross-frequency coupling. The strongest coupling was observed between 1 and 2 Hz phase and 3–20 Hz amplitude, indicating that the phase of delta oscillations (1–3 Hz) modulates the amplitude in the theta (4–8 Hz) and beta (13–30 Hz) bands. The phase of theta oscillations also modulates, to a lesser degree, the amplitude in the beta band (13–30 Hz; Figure S8). We thus observe the same nested hierarchical organization in newborns as in adult humans and monkeys.^{11,12}

To explore if nesting differs across conditions, within the frequency ranges previously identified as relevant (1–8 Hz frequency for phase and 1–30 Hz frequency for amplitude), we submitted the modulation index values from the three language conditions and the two resting state periods to permutation tests with repeated measures ANOVAs with the within-subject factor Condition (InitialRest/French/Spanish/English/FinalRest). Nesting did not differ significantly across conditions in the selected frequency ranges (Figure S9 presents *P*-maps for this comparison).

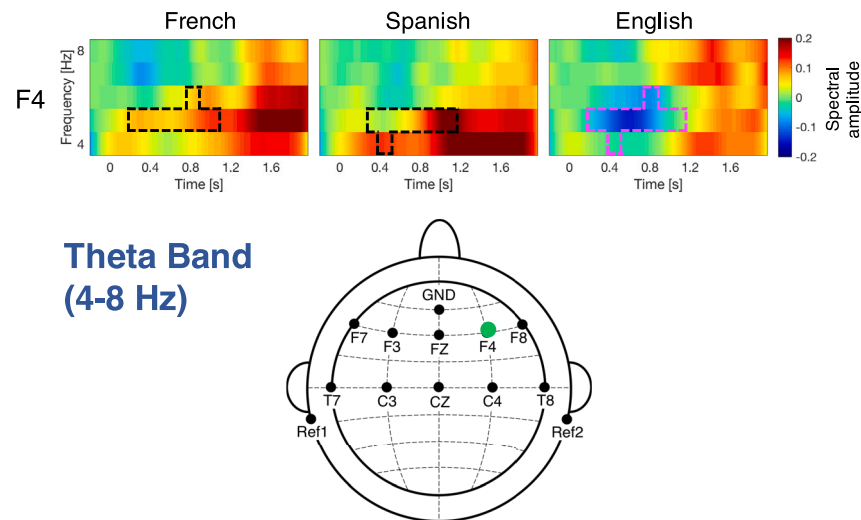


Figure 3. Time-frequency response in the theta band (4–8 Hz) for channels exhibiting significant differences across languages
Channel F4 exhibits higher activation for French and Spanish than for English. Plotting conventions as before. See also [Figure S2](#).

Neural oscillations in the human newborn brain are already hierarchically organized into a nesting relation, in which the phase of the lower-frequency oscillations (mainly delta) modulates the amplitude in the higher-frequency bands (theta and beta), and this organization, present already at rest, is not modified during speech processing.

DISCUSSION

We have explored the developmental origins of neural oscillations in the human brain during rest and speech processing at birth. We recorded newborns' brain activity using electroencephalography in five different conditions: (i) during resting state, prior to auditory stimulation, (ii) while listening to sentences in the prenatally heard language, (iii) in a rhythmically similar unfamiliar language, (iv) in a rhythmically different unfamiliar language, and (v) at rest after auditory stimulation.

When comparing oscillatory activity during speech and pre-stimulus rest, we found that activity in the theta band (and to a lesser extent in the delta band) increased toward the second half of sentence presentation ([Figures 4 and S5](#)). Our findings, therefore, suggest that at birth neural oscillations are already involved in speech processing. The increase in oscillatory activity observed in the delta and theta bands is in line with our prediction that low frequency neural oscillations are active during speech processing, reflecting the processing of sentence prosody. Indeed, in a previous study using the same dataset, we have found that newborns' neural activity in the theta band (more specifically 3–6 Hz) tracks the speech envelope of sentences in the three languages,³⁰ as is also the case in older infants.^{27,31} One interesting possibility is that this may be related to the low-pass filtered nature of the prenatal signal.

Our results do not imply that newborns cannot process phonemes. Indeed, young infants' ability to discriminate phonemic contrasts has been well documented.^{50,51} There is converging evidence, however, that suprasegmental units, such as the syllable or larger prosodic units, may play a particularly important role in infants' early processing and representation of speech.^{52–54} For instance, newborns readily discriminate two words that have different numbers of syllables, but the same number of sub-syllabic units (phonemes or morae), whereas they cannot discriminate words with the same number of syllables, even if they contain different numbers of sub-syllabic units (phonemes or morae).^{55,56} The rhythmic discrimination of languages points in the same direction. While rhythmically similar languages are typically different along many dimensions, newborns fail to discriminate them behaviorally,^{38,39} highlighting the salience of the one property in which they are similar, i.e., rhythm. This salience or privileged status of suprasegmental units, such as the syllable and prosodic units, may have important implications for language acquisition, as they may help the infant segment and parse the input into units that are most relevant to start breaking into language.

In addition to establishing the role of neural oscillations in speech processing at birth, our findings have revealed important differences between the three languages tested. First, as predicted, increased power in the delta ([Figure 2](#)) and theta ([Figure 3](#)) bands was observed consistently for the native language, French, and for the rhythmically similar unfamiliar language, Spanish, as compared to the rhythmically different unfamiliar language, English. Crucially, it is well established that at the behavioral level, newborn infants can discriminate English from French and Spanish and they do so on the basis of rhythmic cues,^{38,39,42} i.e., cues pertaining to the temporal regularities of the speech signal at the syllabic level. Our findings thus reveal, for the first time, the neural basis of newborns' fundamental ability to discriminate rhythmically different languages,^{38,39,41,44} as rhythm is carried by the low-frequency components of the speech signal, specifically the syllabic rate, and might thus be encoded by the low-frequency oscillations delta and theta, as in adults.

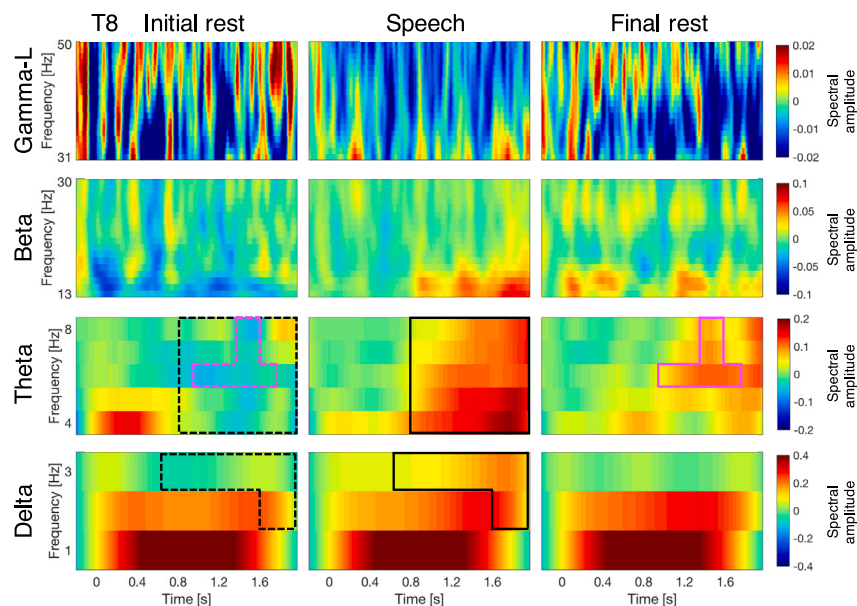


Figure 4. Time-frequency response during resting state (before and after stimulation) and during speech presentation at channel T8

The speech condition corresponds to the average of the three languages (French, Spanish, and English). The time-frequency maps illustrate the mean spectral amplitude per condition. Each frequency band (delta, theta, beta, and low-gamma) is plotted separately, with a different scale, as indicated by the color bars to the right of the time-frequency maps. The black boxes indicate significant differences between initial rest and speech, while magenta boxes indicate significant differences between the two resting state conditions. Boxes with solid lines indicate higher activation, boxes with dashed lines lower activation. See also Figures S5 and S7.

Our results, together with the few existing infant studies on the role of auditory oscillations in language discrimination, suggest that their role may change during the first year of extrauterine life. At birth, discrimination is mainly supported by delta and theta oscillations, at 4.5 months by theta oscillations,²⁵ and from 6 months by gamma oscillations.²⁹ This shows that as the infant brain matures and gains experience with the broadband speech signal, higher frequency oscillations increasingly come into play, perhaps underlying new strategies to discriminate languages not only on the basis of rhythm, but also on the basis of other phonological properties of the native language for the acquisition of which experience is necessary.

Importantly from the point of view of language evolution and language development, the rhythmic discrimination of languages is not a human-specific ability. Monkeys³⁹ and rats⁵⁷ have been shown to be able to perform it as well, suggesting that this ability is not linguistic, but rather acoustic in nature. The cues that operationally define rhythm, i.e., syllable structure, vowels and consonants, have well identifiable and distinguishable acoustic correlates. For instance, vowels have a harmonic spectrum and carry the highest amount of energy in the speech signal, whereas consonants have different non-harmonic, often broad-band spectra, and typically carry less energy. The relative distributions of harmonic, high-energy segments and broad-band, transient, lower energy segments thus offers a reliable acoustic cue to discriminate between rhythmically different languages. It is thus not surprising that non-linguistic animals such as monkeys can also perform the discrimination. It may even be possible that they use similar oscillatory mechanisms to achieve it, as hierarchically embedded neural oscillations similar to those in humans have also been observed in monkeys' auditory cortex.¹² From the point of view of language evolution, this suggests that language may have recruited already available auditory mechanisms, and more relevantly for our study, from the perspective of language development, it suggests that such an early and fundamental task as discriminating and identifying the target language(s) relies on basic auditory mechanisms, as language experience and linguistic processing are still limited early in life.

Second, the differential activation favoring French and Spanish may also be related to the familiarity of the rhythmicity of these languages, i.e., prenatal experience. This cannot be established conclusively, as the results may also reflect responses triggered by the specific acoustic properties of French and Spanish as opposed to English independently of their familiarity. Future research testing the same stimuli with prenatally English-exposed newborns will provide a definitive answer.

Importantly, these differences in oscillatory activity across languages do not relate to the neonate brain's ability to track the speech envelope. In a previous study³⁰ using the same dataset, we found that newborns' brain responses track the envelope of the speech signal both in phase (in the 3–6 Hz range) and in amplitude (in the 1–40 Hz range) regardless of language familiarity. Our previous and current findings taken together suggest that the ability to discriminate rhythmically different languages at birth requires the simultaneous processing of time and frequency information along multiple scales.

In an attempt to understand how language experience may shape the brain, we assessed whether stimulation after-effects are already present at birth by comparing oscillatory activity from two resting state periods: before and after newborns had been exposed to a relatively

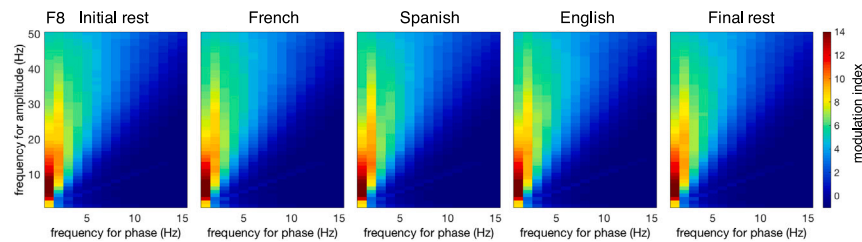


Figure 5. The modulation index as a function of analytic phase (1–15 Hz) and analytic amplitude (1–50 Hz) for each condition at channel F8
Larger values indicate stronger cross-frequency coupling. The color bar to the right of the figure shows the modulation index scale for all plots. See also [Figure S8](#).

long period (21 min) of speech stimulation. Our results reveal that power in the theta band was higher during the final resting state period than at the initial one ([Figures 4 and S7](#)). These differences are similar to those observed between activation during the initial resting state period and speech. These findings converge with results for 3-month-old infants tested with near-infrared spectroscopy.⁴⁶ Further studies are necessary to investigate the exact functional role of this post-stimulation effect, but it is compatible with the interpretation often given in the developmental literature⁴⁶ that it may support learning and perceptual attunement to the speech stimuli encountered.

An important property of neural oscillations, instrumental to their ability to support speech processing, is their hierarchical organization. This mirrors the hierarchical organization of speech⁵⁸ and allows oscillations to be linked, so information about smaller units is processed within information about larger units. For oscillations to be useful for speech processing and language acquisition from the start, it is crucial to establish whether they are hierarchically organized at birth. Our results indeed uncover such a hierarchy in newborns, with maximal coupling between the phase of delta oscillations and the amplitude of theta and beta oscillations ([Figures 5 and S8](#)). These findings mesh well with the nesting relations found between the phase of lower-frequency oscillations and the amplitude of higher-frequency oscillations in adults (theta and gamma coupling^{11,49}), and animals (delta and theta coupling¹²). While our newborn results show a nesting relation between a broader range of frequencies, likely due to the overall slower and less structured nature of newborn EEG,^{59–61} the direction of the hierarchy is preserved. Hierarchical organization may thus start broad and unspecialized at birth, and may get fine-tuned through development. This suggestion is supported by the fact that the nesting relationship is not modulated by speech stimulation as compared to rest or by the language heard, which could indicate that it requires extended experience to fully develop.

From a more general theoretical perspective, we may ask at what level newborns process speech stimuli. After all, monkeys are also able to discriminate languages based on rhythm³⁹ and exhibit embedded neural oscillations in response to auditory stimuli.¹² Furthermore, the human brain entrains to non-linguistic auditory stimuli, as well.^{33–37} It is thus possible that newborns process speech at the auditory level or already at a more language-specific level. Our study was not designed to adjudicate between the two possibilities. We note, however, that processing doesn't need to be speech-specific for it to be useful for subsequent language development. Indeed, phonological bootstrapping theories of language acquisition⁶² posit that infants may exploit correlations between sound patterns in speech at the most basic acoustic dimensions of pitch, intensity or duration, and more abstract lexical or grammatical properties to break into language and bootstrap these more abstract structures.

Our study targeted a small number of focal electrode sites known to be involved in auditory and speech processing in newborns⁶³ in order to increase infants' comfort and the feasibility of the study. Future studies with high density montages could be more informative about the spatial localizations of the observed effects, as they may allow source-space reconstruction (albeit this remains challenging for developmental EEG).

In summary, we have shown, in a large sample of newborn infants, that an organized hierarchy of neural oscillations is already present at the start of extrauterine experience with language, and supports key speech perception abilities such as the rhythmic discrimination of languages. This neural architecture lays the foundations for subsequent language development, explaining how infants acquire language so rapidly and effortlessly.

Limitations of the study

Here we have provided new evidence on the neural mechanisms underlying speech processing and language discrimination at birth. Our work however has some clear limitations. *The first limitation* pertains to testing time constraints. When designing infant experiments testing sessions need to be kept short. Due to this, the resting state blocks were shorter than the language blocks. Moreover, this time constraint prevented us from including additional conditions, which leads us to our *second limitation*. In the current study we could not test non-linguistic auditory stimuli, such as music or environmental sounds. Future studies will require testing such conditions to better understand how neural oscillations support auditory processing generally as compared to speech processing at birth.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- [KEY RESOURCES TABLE](#)
- [RESOURCE AVAILABILITY](#)

- Lead contact
- Materials availability
- Data and code availability
- **EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS**
 - Participants
- **METHOD DETAILS**
 - Procedure
 - Stimuli
 - Data acquisition
 - EEG processing
- **QUANTIFICATION AND STATISTICAL ANALYSIS**
 - Time-frequency analysis
 - Nesting analysis

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.isci.2023.108187>.

ACKNOWLEDGMENTS

This work was supported by the “PredictAble” Marie Skłodowska-Curie Action (MSCA) Innovative Training Network (ITN) grant, the ECOSud action nr. C20S02, the ERC Consolidator Grant 773202 “BabyRhythm”, the ANR’s French Investissements d’Avenir – Labex EFL Program under Grant [ANR-10-LABX-0083], the Italian Ministry for Universities and Research FARE grant nr. R204MPRHKE, the European Union Next Generation EU NRRP M6C2 - Investment 2.1 - SYNPHONIA Project, as well as the Italian Ministry for Universities and Research PRIN grant nr. 2022WX3FM5 to Judit Gervain. We would like to thank the Robert Debré Hospital for providing access to the newborns, and all the families and their babies for their participation in this study. We sincerely acknowledge Lucie Martin and Anouche Banikyan for their help with infant testing; as well as Gábor Háden and Eugenio Parise for valuable discussions on data analysis.

AUTHOR CONTRIBUTIONS

M.O. and J.G. conceived and designed the experiment. M.O. performed the experiment. M.O. and R.G. analyzed the data. M.O., R.G. and J.G. wrote the article. All the authors evaluated the results and edited the article.

DECLARATION OF INTERESTS

The authors declare no competing interests.

INCLUSION AND DIVERSITY

We support inclusive, diverse, and equitable conduct of research.

Received: September 6, 2022

Revised: August 29, 2023

Accepted: October 9, 2023

Published: October 12, 2023

REFERENCES

1. Skeide, M.A., and Friederici, A.D. (2016). The ontogeny of the cortical language network. *Nat. Rev. Neurosci.* 17, 323–332. <https://doi.org/10.1038/nrn.2016.23>.
2. Bastiaansen, M., and Hagoort, P. (2006). Oscillatory neuronal dynamics during language comprehension. *Prog. Brain Res.* 159, 179–196. [https://doi.org/10.1016/S0079-6123\(06\)59012-0](https://doi.org/10.1016/S0079-6123(06)59012-0).
3. Giraud, A.-L., and Poeppel, D. (2012). Cortical oscillations and speech processing: emerging computational principles and operations. *Nat. Neurosci.* 15, 511–517. <https://doi.org/10.1038/nn.3063>.
4. Doelling, K.B., Arnal, L.H., Ghitza, O., and Poeppel, D. (2014). Acoustic landmarks drive delta–theta oscillations to enable speech comprehension by facilitating perceptual parsing. *Neuroimage* 85 Pt 2, 761–768. <https://doi.org/10.1016/j.neuroimage.2013.06.035>.
5. Gross, J., Hoogenboom, N., Thut, G., Schyns, P., Panzeri, S., Belin, P., and Garrod, S. (2013). Speech Rhythms and Multiplexed Oscillatory Sensory Coding in the Human Brain. Poeppel D, editor. *PLoS Biol.* 11, e1001752. <https://doi.org/10.1371/journal.pbio.1001752>.
6. Meyer, L. (2018). The neural oscillations of speech processing and language comprehension: state of the art and emerging mechanisms. *Eur. J. Neurosci.* 48, 2609–2621. <https://doi.org/10.1111/ejn.13748>.
7. Arnal, L.H., Poeppel, D., and Giraud, A. (2015). Temporal coding in the auditory cortex. *Handb. Clin. Neurol.* 129, 85–98. <https://doi.org/10.1016/B978-0-444-62630-1.00005-6>.
8. Giraud, A.-L., Kleinschmidt, A., Poeppel, D., Lund, T.E., Frackowiak, R.S.J., and Laufs, H. (2007). Endogenous Cortical Rhythms Determine Cerebral Specialization for Speech Perception and Production. *Neuron* 56, 1127–1134. <https://doi.org/10.1016/j.neuron.2007.09.038>.
9. Morillon, B., Lehongre, K., Frackowiak, R.S.J., Ducorps, A., Kleinschmidt, A., Poeppel, D., and Giraud, A.-L. (2010). Neurophysiological origin of human brain asymmetry for speech and language. *Proc. Natl. Acad. Sci. USA* 107,

- 18688–18693. <https://doi.org/10.1073/pnas.1007189107>.
10. Ding, N., Melloni, L., Zhang, H., Tian, X., and Poeppel, D. (2016). Cortical tracking of hierarchical linguistic structures in connected speech. *Nat. Neurosci.* 19, 158–164. <https://doi.org/10.1038/nn.4186>.
 11. Morillon, B., Liégeois-Chauvel, C., Arnal, L.H., Bénar, C.G., and Giraud, A.-L. (2012). Asymmetric Function of Theta and Gamma Activity in Syllable Processing: An Intra-Cortical Study. *Front. Psychol.* 3, 248. <https://doi.org/10.3389/fpsyg.2012.00248>.
 12. Lakatos, P., Shah, A.S., Knuth, K.H., Ulbert, I., Karmos, G., and Schroeder, C.E. (2005). An Oscillatory Hierarchy Controlling Neuronal Excitability and Stimulus Processing in the Auditory Cortex. *J. Neurophysiol.* 94, 1904–1911. <https://doi.org/10.1152/jn.00263.2005>.
 13. Steriade, M., Contreras, D., Amzica, F., and Timofeev, I. (1996). Synchronization of fast (30–40 Hz) spontaneous oscillations in intrathalamic and thalamocortical networks. *J. Neurosci.* 16, 2788–2808. <https://doi.org/10.1523/JNEUROSCI.16-08-02788.1996>.
 14. Lakatos, P., Szilágyi, N., Pincze, Z., Rajkai, C., Ulbert, I., and Karmos, G. (2004). Attention and arousal related modulation of spontaneous gamma-activity in the auditory cortex of the cat. *Brain Res. Cogn. Brain Res.* 19, 1–9. <https://doi.org/10.1016/j.cogbrainres.2003.10.023>.
 15. Di Liberto, G.M., O’Sullivan, J.A., and Lalor, E.C. (2015). Low-Frequency Cortical Entrainment to Speech Reflects Phoneme-Level Processing. *Curr. Biol.* 25, 2457–2465. <https://doi.org/10.1016/j.cub.2015.08.030>.
 16. Peelle, J.E., Gross, J., and Davis, M.H. (2013). Phase-Locked Responses to Speech in Human Auditory Cortex are Enhanced During Comprehension. *Cerebr. Cortex* 23, 1378–1387. <https://doi.org/10.1093/cercor/bhs118>.
 17. Bourguignon, M., De Tiege, X., De Beeck, M.O., Ligot, N., Paquier, P., Van Bogaert, P., Goldman, S., Hari, R., and Jousmäki, V. (2013). The pace of prosodic phrasing couples the listener’s cortex to the reader’s voice. *Hum. Brain Mapp.* 34, 314–326. <https://doi.org/10.1002/hbm.21442>.
 18. Ahissar, E., Nagarajan, S., Ahissar, M., Protopapas, A., Mahncke, H., and Merzenich, M.M. (2001). Speech comprehension is correlated with temporal response patterns recorded from auditory cortex. *Proc. Natl. Acad. Sci. USA* 98, 13367–13372. <https://doi.org/10.1073/pnas.201400998>.
 19. Nourski, K.V., Reale, R.A., Oya, H., Kawasaki, H., Kovach, C.K., Chen, H., Howard, M.A., and Brugge, J.F. (2009). Temporal Envelope of Time-Compressed Speech Represented in the Human Auditory Cortex. *J. Neurosci.* 29, 15564–15574. <https://doi.org/10.1523/JNEUROSCI.3065-09.2009>.
 20. Neuling, T., Rach, S., Wagner, S., Wolters, C.H., and Herrmann, C.S. (2012). Good vibrations: Oscillatory phase shapes perception. *Neuroimage* 63, 771–778. <https://doi.org/10.1016/j.neuroimage.2012.07.024>.
 21. Riecke, L., Formisano, E., Sorger, B., Başkent, D., and Gaudrain, E. (2018). Neural Entrainment to Speech Modulates Speech Intelligibility. *Curr. Biol.* 28, 161–169.e5. <https://doi.org/10.1016/j.cub.2017.11.033>.
 22. Bastiaansen, M., Magyari, L., and Hagoort, P. (2010). Syntactic Unification Operations Are Reflected in Oscillatory Dynamics during On-line Sentence Comprehension. *J. Cognit. Neurosci.* 22, 1333–1347. <https://doi.org/10.1162/jocn.2009.21283>.
 23. Pefkou, M., Arnal, L.H., Fontolan, L., and Giraud, A.-L. (2017). θ -Band and β -Band Neural Activity Reflects Independent Syllable Tracking and Comprehension of Time-Compressed Speech. *J. Neurosci.* 37, 7930–7938. <https://doi.org/10.1523/JNEUROSCI.2882-16.2017>.
 24. Bosseler, A.N., Taulu, S., Pihko, E., Mäkelä, J.P., Imada, T., Ahonen, A., and Kuhl, P.K. (2013). Theta brain rhythms index perceptual narrowing in infant speech perception. *Front. Psychol.* 4, 690. <https://doi.org/10.3389/fpsyg.2013.00690>.
 25. Nacar Garcia, L., Guerrero-Mosquera, C., Colomer, M., and Sebastian-Galles, N. (2018). Evoked and oscillatory EEG activity differentiates language discrimination in young monolingual and bilingual infants. *Sci. Rep.* 8, 2770. <https://doi.org/10.1038/s41598-018-20824-0>.
 26. Parise, E., and Csibra, G. (2013). Neural Responses to Multimodal Ostensive Signals in 5-Month-Old Infants. *Becchio C, editor. PLoS One* 8, e72360. <https://doi.org/10.1371/journal.pone.0072360>.
 27. Ortiz-Mantilla, S., Hämäläinen, J.A., Musacchia, G., and Benasich, A.A. (2013). Enhancement of Gamma Oscillations Indicates Preferential Processing of Native over Foreign Phonemic Contrasts in Infants. *J. Neurosci.* 33, 18746–18754. <https://doi.org/10.1523/JNEUROSCI.3260-13.2013>.
 28. Ortiz-Mantilla, S., Hämäläinen, J.A., Realpe-Bonilla, T., and Benasich, A.A. (2016). Oscillatory Dynamics Underlying Perceptual Narrowing of Native Phoneme Mapping from 6 to 12 Months of Age. *J. Neurosci.* 36, 12095–12105. <https://doi.org/10.1523/JNEUROSCI.1162-16.2016>.
 29. Peña, M., Pittaluga, E., and Mehler, J. (2010). Language acquisition in premature and full-term infants. *Proc. Natl. Acad. Sci. USA* 107, 3823–3828. <https://doi.org/10.1073/pnas.0914326107>.
 30. Ortiz Barajas, M.C., Guevara, R., and Gervain, J. (2021). The origins and development of speech envelope tracking during the first months of life. *Dev. Cogn. Neurosci.* 48, 100915. <https://doi.org/10.1016/j.dcn.2021.100915>.
 31. Kalashnikova, M., Peter, V., Di Liberto, G.M., Lalor, E.C., and Burnham, D. (2018). Infant-directed speech facilitates seven-month-old infants’ cortical tracking of speech. *Sci. Rep.* 8, 13745. <https://doi.org/10.1038/s41598-018-32150-6>.
 32. Goswami, U. (2019). Speech rhythm and language acquisition: an amplitude modulation phase hierarchy perspective. *Ann. N. Y. Acad. Sci.* 1453, 67–78. <https://doi.org/10.1111/nyas.14137>.
 33. Nozaradan, S., Peretz, I., Missal, M., and Mouraux, A. (2011). Tagging the Neuronal Entrainment to Beat and Meter. *J. Neurosci.* 31, 10234–10240. <https://doi.org/10.1523/JNEUROSCI.0411-11.2011>.
 34. Cirelli, L.K., Spinelli, C., Nozaradan, S., and Trainor, L.J. (2016). Measuring Neural Entrainment to Beat and Meter in Infants: Effects of Music Background. *Front. Neurosci.* 10, 229. <https://doi.org/10.3389/fnins.2016.00229>.
 35. Fujioka, T., Trainor, L.J., Large, E.W., and Ross, B. (2009). Beta and Gamma Rhythms in Human Auditory Cortex during Musical Beat Processing. *Ann. N. Y. Acad. Sci.* 1169, 89–92. <https://doi.org/10.1111/j.1749-6632.2009.04779.x>.
 36. Fujioka, T., Trainor, L.J., Large, E.W., and Ross, B. (2012). Internalized Timing of Isochronous Sounds Is Represented in Neuromagnetic Beta Oscillations. *J. Neurosci.* 32, 1791–1802. <https://doi.org/10.1523/JNEUROSCI.4107-11.2012>.
 37. Cirelli, L.K., Bosnyak, D., Manning, F.C., Spinelli, C., Marie, C., Fujioka, T., Ghahremani, A., and Trainor, L.J. (2014). Beat-induced fluctuations in auditory cortical beta-band activity: using EEG to measure age-related changes. *Front. Psychol.* 5, 742. <https://doi.org/10.3389/fpsyg.2014.00742>.
 38. Nazzi, T., Bertoncini, J., and Mehler, J. (1998). Language discrimination by newborns: Toward an understanding of the role of rhythm. *J. Exp. Psychol. Hum. Percept. Perform.* 24 (3), 756–766. <https://doi.org/10.1037/0096-1523.24.3.756>.
 39. Ramus, F., Hauser, M.D., Miller, C., Morris, D., and Mehler, J. (2000). Language Discrimination by Human Newborns and by Cotton-Top Tamarin Monkeys. *Science* 288, 349–351. <https://doi.org/10.1126/science.288.5464.349>.
 40. Ramus, F., Nespor, M., and Mehler, J. (1999). Correlates of linguistic rhythm in the speech signal. *Cognition* 73, 265–292. [https://doi.org/10.1016/S0010-0277\(99\)00058-X](https://doi.org/10.1016/S0010-0277(99)00058-X).
 41. Mehler, J., Jusczyk, P., Lambert, G., Halsted, N., Bertoncini, J., and Amiel-Tison, C. (1988). A precursor of language acquisition in young infants. *Cognition* 29, 143–178. [https://doi.org/10.1016/0010-0277\(88\)90035-2](https://doi.org/10.1016/0010-0277(88)90035-2).
 42. Byers-Heinlein, K., Burns, T.C., and Werker, J.F. (2010). The Roots of Bilingualism in Newborns. *Psychol. Sci.* 21, 343–348. <https://doi.org/10.1177/0956797609360758>.
 43. Pujol, R., Lavigne-rebillard, M., and Uziel, A. (1991). Development of the Human Cochlea. *Acta Otolaryngol. Suppl.* 111, 7–13. <https://doi.org/10.3109/00016489109128023>.
 44. Moon, C., Cooper, R.P., and Fifer, W.P. (1993). Two-day-olds prefer their native language. *Infant Behav. Dev.* 16, 495–500. [https://doi.org/10.1016/0163-6383\(93\)80007-U](https://doi.org/10.1016/0163-6383(93)80007-U).
 45. Mampe, B., Friederici, A.D., Christophe, A., and Wermke, K. (2009). Newborns’ Cry Melody Is Shaped by Their Native Language. *Curr. Biol.* 19, 1994–1997. <https://doi.org/10.1016/j.cub.2009.09.064>.
 46. Homae, F., Watanabe, H., Nakano, T., and Taga, G. (2011). Large-Scale Brain Networks Underlying Language Acquisition in Early Infancy. *Front. Psychol.* 2, 93. <https://doi.org/10.3389/fpsyg.2011.00093>.
 47. Werker, J.F. (2018). Perceptual beginnings to language acquisition. *Appl. Psycholinguist.* 39, 703–728. <https://doi.org/10.1017/S0142716418000152>.
 48. Ortiz Barajas, M.C., and Gervain, J. (2021). The Role of Prenatal Experience and Basic Auditory Mechanisms in the Development of Language. In *Minnesota Symposia on Child Psychology*, 1st ed., M.D. Sera and M. Koenig, eds. (Wiley), pp. 88–112. <https://doi.org/10.1002/9781119684527.ch4>.
 49. Canolty, R.T., Edwards, E., Dalal, S.S., Soltani, M., Nagarajan, S.S., Kirsch, H.E., Berger, M.S., Barbaro, N.M., and Knight, R.T. (2006). High Gamma Power Is Phase-Locked to Theta Oscillations in Human Neocortex. *Science* 313, 1626–1628. <https://doi.org/10.1126/science.1128115>.
 50. Eimas, P.D., Siqueland, E.R., Jusczyk, P., and Vigorito, J. (1971). Speech Perception in

- Infants. *Science* 171, 303–306. <https://doi.org/10.1126/science.171.3968.303>.
51. Dehaene-Lambertz, G., and Baillet, S. (1998). A phonological representation in the infant brain. *Neuroreport* 9, 1885–1888. <https://doi.org/10.1097/00001756-199806010-00040>.
 52. Mehler, J., and Bertoncini, J. (1979). Infants' perception of speech and other acoustic stimuli. In *Psycholinguistics 2: Structure and Process* (MIT Press), pp. 69–105.
 53. Mehler, J., Dommergues, J.Y., Frauenfelder, U., and Segui, J. (1981). The syllable's role in speech segmentation. *J. Verb. Learn. Verb. Behav.* 20, 298–305. [https://doi.org/10.1016/S0022-5371\(81\)90450-3](https://doi.org/10.1016/S0022-5371(81)90450-3).
 54. Gervain, J. (2015). Plasticity in early language acquisition: the effects of prenatal and early childhood experience. *Curr. Opin. Neurobiol.* 35, 13–20. <https://doi.org/10.1016/j.conb.2015.05.004>.
 55. Bijeljac-Babic, R., Bertoncini, J., and Mehler, J. (1993). How do 4-day-old infants categorize multisyllabic utterances? *Dev. Psychol.* 29, 711–721. <https://doi.org/10.1037/0012-1649.29.4.711>.
 56. Bertoncini, J., Floccia, C., Nazzi, T., and Mehler, J. (1995). Morae and Syllables: Rhythmical Basis of Speech Representations in Neonates. *Lang. Speech* 38, 311–329. <https://doi.org/10.1177/002383099503800401>.
 57. Toro, J.M., Trobalon, J.B., and Sebastián-Gallés, N. (2003). The use of prosodic cues in language discrimination tasks by rats. *Anim. Cognit.* 6, 131–136. <https://doi.org/10.1007/s10071-003-0172-0>.
 58. Nespor, M., and Vogel, I. (1988). *Prosodic phonology*. Dordrecht: Foris Publications. Pp. xiv + 327. *Phonology* 5, 161–168. <https://doi.org/10.1017/S0952675700002219>.
 59. Torres, F., and Anderson, C. (1985). The Normal EEG of the Human Newborn. *J. Clin. Neurophysiol.* 2, 89–103. <https://doi.org/10.1097/00004691-198504000-00001>.
 60. De Haan, M. (2007). Infant EEG and Event-Related Potentials (Psychology Press). <https://doi.org/10.4324/9780203759660>.
 61. Britton, J.W., Frey, L.C., Hopp, J.L., Korb, P., Koubeissi, M.Z., Lievens, W.E., Pestana-Knight, E.M., and St. Louis, E.K. (2016). In *Electroencephalography (EEG): An Introductory Text and Atlas of Normal and Abnormal Findings in Adults, Children, S. Infants, E.K. Louis, and L.C. Frey, eds.* (American Epilepsy Society). <https://doi.org/10.5698/978-0-9979756-0-4>.
 62. Morgan, J.L., and Demuth, K. (1996). *Signal to Syntax: Bootstrapping from Speech to Grammar in Early Acquisition* (Psychology Press).
 63. Tóth, B., Urbán, G., Háden, G.P., Márk, M., Török, M., Stam, C.J., and Winkler, I. (2017). Large-scale network organization of EEG functional connectivity in newborn infants. *Hum. Brain Mapp.* 38, 4019–4033. <https://doi.org/10.1002/hbm.23645>.
 64. 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, 1976–1989. <https://doi.org/10.1121/1.5006179>.
 65. Stefanics, G., Háden, G.P., Sziller, I., Balázs, L., Beke, A., and Winkler, I. (2009). Newborn infants process pitch intervals. *Clin. Neurophysiol.* 120, 304–308. <https://doi.org/10.1016/j.clinph.2008.11.020>.
 66. Csibra, G., Davis, G., Spratling, M.W., and Johnson, M.H. (2000). Gamma Oscillations and Object Processing in the Infant Brain. *Science* 290, 1582–1585. <https://doi.org/10.1126/science.290.5496.1582>.
 67. Maris, E., and Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *J. Neurosci. Methods* 164, 177–190. <https://doi.org/10.1016/j.jneumeth.2007.03.024>.
 68. Hülsemann, M.J., Naumann, E., and Rasch, B. (2019). Quantification of Phase-Amplitude Coupling in Neuronal Oscillations: Comparison of Phase-Locking Value, Mean Vector Length, Modulation Index, and Generalized-Linear-Modeling-Cross-Frequency-Coupling. *Front. Neurosci.* 13, 573. <https://doi.org/10.3389/fnins.2019.00573>.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Software and algorithms		
MATLAB_R2018b	MathWorks	https://fr.mathworks.com/products/matlab.html
WTools	Parise and Csibra, 2013 ²⁶	N/A

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources should be directed to and will be fulfilled by the lead contact, Maria Clemencia Ortiz-Barajas (maria.c.ortizb@gmail.com).

Materials availability

This study did not generate new reagents.

Data and code availability

- The EEG data reported in this study cannot be deposited in a public repository because it represents sensitive medical information. The data are available from the [lead contact](#) upon reasonable request.
- This paper does not report original code.
- Any additional information required to reanalyze the data reported in this paper is available from the [lead contact](#) upon request.

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

The EEG data from this study was acquired as part of a larger project that aimed to investigate speech perception and processing during the first two years of life. One previous publication presented a superset of the current dataset (47 participants) evaluating speech envelope tracking in newborns and 6-month-olds.³⁰ The dataset used in this manuscript (40 participants) represents a subset of that used in our previous publication, as 7 participants were rejected due to high frequency noise.

Participants

The protocol for this study was approved by the CER Paris Descartes ethics committee of the Paris Descartes University (currently, Université Paris Cité). All parents gave written informed consent prior to participation, and were present during the testing session. We recruited participants at the maternity ward of the Robert-Debré Hospital in Paris, and we tested them during their hospital stay. 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. We tested a total of 54 newborns, and excluded 14 participants from data analysis due to: not finishing the experiment due to fussiness and crying (n = 2), technical problems (n = 1), or bad data quality resulting in an insufficient number of non-rejected trials in the language conditions (n = 11). Thus, electrophysiological data from 40 newborns (age 2.55 ± 1.24 days; range 1–5 days; 17 girls, 23 boys) were included in the analyses of the language conditions (French, Spanish, English). Some of these 40 participants were excluded from the analyses of the resting state conditions due to bad data quality resulting in an insufficient number of non-rejected trials in these conditions (n = 4); therefore, EEG data from 36 newborns were included in the analyses comparing the language conditions to the initial resting state condition, and in the analyses comparing the two resting state conditions (initial vs. final).

METHOD DETAILS

Procedure

We tested infants by recording their neural activity using electroencephalography (EEG) during resting state, and while presenting them with naturally spoken sentences in three languages. 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 ([Table S1](#)): 11 newborns heard set1, 15 newborns heard set2, and 14 newborns heard set3. During the recording session, newborns were comfortably asleep or at rest in their hospital bassinets ([Figure 1A](#)). 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). We presented participants with one sentence per language, and repeated it 100 times to ensure sufficiently good data quality. The experiment consisted of 5 blocks: one

initial resting state block, three language blocks, and one final resting state block (Figure 1C). Each resting state block lasted 3 min, while each language block, containing the 100 repetitions of the test sentence of the given language, lasted around 7 min. An interstimulus interval of random duration (between 1 and 1.5 s) was introduced between sentence repetitions, and an interblock interval of 10 s was introduced between language blocks (Figure 1C). The order of the languages was pseudo-randomized and approximately counterbalanced across participants. The entire recording session lasted about 27 min.

The reason for a shorter resting state recording was purely practical, related to the time constraints for testing newborns at the maternity: factors such as medical check-ups, baby feedings, mother's fatigue, and visitors, reduce the families' availability to participate. A newborn experiment needs to be as short as possible to increase its feasibility. Importantly, this asymmetry in recording times (3 min for resting state blocks vs. 7 min for language blocks) does not translate into a similarly large asymmetry in the number of segmented epochs, as we obtained 100 epochs per language block and 70 epochs per resting state block.

Stimuli

We tested infants at rest and during speech stimulation. While at rest, no stimulus was presented, and during speech stimulation, we tested them 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 by adding or removing adjectives (or phrases) from certain sentences in order to match the duration and syllable count across languages within the same set (see Table S1). 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. There were no significant differences between the sentences in the three languages in terms of minimum and maximum pitch, pitch range and average pitch (see Table S1). The audio files (.wav) of the 9 utterances used as stimuli are included as part of the supplementary information (Audios S1–S9). Figure S10 displays the sentences' time-series, and Figure S11 displays their frequency spectra, obtained using the Fast Fourier Transform function in MATLAB (fft). We computed the amplitude and frequency modulation spectra of the sentences in the three languages as defined by Varnet and colleagues⁶⁴ to explore if utterances were consistently different across languages (Figure S12). We found that utterances were similar in every spectral decomposition. The intensity of all recordings was adjusted to 77 dB.

Data acquisition

We recorded the EEG data with active electrodes and an acquisition system from Brain Products (actiCAP & actiCHamp, Brain Products GmbH, Gilching, Germany). We used a 10-channel layout to acquire cortical responses from the following scalp positions: F7, F3, FZ, F4, F8, T7, C3, CZ, C4, T8 (Figure 1B). We chose these recording locations in order to include those where auditory and speech perception related neural responses are typically observed in infants^{63,65} (channels T7 and T8 used to be called T3 and T4 respectively). We used two additional electrodes placed on each mastoid for online reference, and a ground electrode placed on the forehead. Data were referenced online to the average of the two mastoid channels, and they were not re-referenced offline. Data were recorded at a sampling rate of 500 Hz, and online filtered with a high cutoff filter at 200 Hz, a low cutoff filter at 0.01 Hz and an 8 kHz (–3 dB) anti-aliasing filter. The electrode impedances were kept below 140 k Ω .

EEG processing

We processed the EEG data using custom MATLAB scripts. 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 50 Hz using a zero phase-shift Chebyshev filter, and segmented them into a series of 2,560-ms long epochs. Each epoch started 400 ms before the utterance onset (corresponding to the pre-stimulus baseline), and contained a 2,160 ms long post-stimulus interval (corresponding to the duration of the shortest sentence). For the resting state data (3 min recorded during silence at the beginning of the experiment and 3 min at the end), we followed the same epoching procedure described above, arbitrarily segmenting the 3 min into 2,560-ms long epochs, which yielded 70 epochs for each rest condition. All the processing steps that follow were applied to the language conditions (French, Spanish, English) and the resting state conditions (initial and final rest) in a similar fashion. We submitted all epochs to a three-stage rejection process to exclude the contaminated ones. First, we rejected epochs with peak-to-peak amplitude exceeding 150 μ V. Second, we rejected those whose standard deviation (SD) was higher than 3 times the mean SD of all non-rejected epochs, or lower than one-third the mean SD. Third, we visually inspected the remaining epochs to remove any residual artifacts. Participants who had less than 20 remaining epochs in a given condition after epoch rejection were excluded from the data analysis involving that condition ($n = 11$ for the language conditions, $n = 2$ for the initial rest, and $n = 2$ for the final rest, as also reported in the participants section). Non-excluded participants contributed on average 43 epochs (SD: 14.3; range across participants: 20–83) for French, 41 epochs (SD: 13.4; range across participants: 20–70) for Spanish, 37 epochs (SD: 12.7; range across participants: 20–70) for English, 38 epochs (SD: 10.0; range across participants: 20–57) for the initial rest, and 43 epochs (SD: 11.8; range across participants: 20–65) for the final rest. We submitted the number of non-rejected epochs from all the participants to a repeated measures ANOVA with the within-subject factor Condition (InitialRest/French/Spanish/English/FinalRest), and it yielded no significant main effect of Condition ($F(4) = 1.7368$, $p = 0.1452$).

QUANTIFICATION AND STATISTICAL ANALYSIS

Time-frequency analysis

We subjected the non-rejected epochs to time-frequency analysis to uncover stimulus-evoked oscillatory responses using the MATLAB toolbox 'WTools'.²⁶ With this toolbox we performed a continuous wavelet transform of each non-rejected epoch using Morlet wavelets (number of cycles 3.5) at 1 Hz intervals in the 1–50 Hz range. The full pipeline is described in detail in.^{26,66} Briefly, complex Morlet Wavelets are computed at steps of 1 Hz with a sigma of 3.5. The real and the imaginary parts of the wavelets are computed separately as cos and sin components, respectively. The signal is then convoluted with each wavelet. The absolute value of each complex coefficient is then computed. Signal edges are chopped to avoid distortions. Finally, a subtractive baseline correction is performed for each frequency.

Time-frequency transformed epochs were then averaged for each condition separately. To remove the distortion introduced by the wavelet transform, the first and last 200 ms of the epochs were removed, resulting in 2,160 ms long segments, including 200 ms before and 1,960 ms after stimulus onset. The averaged epochs were then baseline corrected using the mean amplitude of the 200 ms pre-stimulus window as baseline, subtracting it from the whole epoch at each frequency. This process resulted in a time-frequency map of spectral amplitude values (not power) per condition and channel, at the participant level. The group mean of these time-frequency maps for channel F8 is presented in [Figure 1D](#) as an example ([Figure S1](#) for other channels). To explore differences in the time-frequency responses across conditions, we submitted the time-frequency maps to various statistical analyses.

Language differences

To assess whether oscillatory activity differs across the three language conditions we submitted the spectral amplitude values from their time-frequency responses to repeated measures ANOVAs with the within-subject factor Language (French/Spanish/English) (see [Figure S2A](#) for the *P*-maps for this analysis). We calculated cluster-level statistics and performed nonparametric statistical testing by calculating the *p* value of the clusters under the permutation distribution,⁶⁷ which was obtained by permuting the language labels in the original dataset 1000 times (see [Figure S2C](#) for examples of the permutation distributions). As post hoc analyses, we computed paired samples *t*-tests (two-tailed) to compare the languages pairwise, which helped us reveal the direction of the language effects. The sample size for these analyses was 40 participants.

Additionally, we performed the same language comparison in a subset of the data (14 participants who listened to the sentences in Set 3). For these analyses we performed a two-stage rejection process (the peak-to-peak amplitude rejection step, and the standard deviation step) as before, but not the visual inspection stage, due to time constraints. We executed the same statistical analyses as before (repeated measures ANOVAs, and cluster-level statistics). We obtained the permutation distribution by permuting the language labels 100 times. See [Figures S3](#) and [S4](#) for the results of these analyses.

Resting state vs. speech processing

To investigate whether neural oscillations are modulated by speech at birth, we explored differences between the oscillatory activity during resting state and during speech processing. In order to evaluate speech processing in a more general, non-language specific way, we averaged the time-frequency responses from the three languages (French, Spanish and English) to obtain a speech stimulation response. We submitted the spectral amplitude values from the time-frequency responses from the initial resting state condition and the speech stimulation condition to paired-samples *t*-tests (two tailed) (see [Figure S5A](#) for the *P*-maps for this analysis). Cluster-level statistics and permutation distributions were computed by permuting the conditions' labels in the original dataset 1000 times, as before. The sample size for these analyses was 36 participants. Additionally, we computed the differences between the oscillatory activity during the initial resting period and during the processing of each language separately for reference ([Figure S6](#)).

Stimulation after-effects

To investigate whether speech stimulation has a lasting effect on resting state oscillations, we compared the time-frequency responses from the initial and final resting state periods by submitting their spectral amplitude values to paired-samples *t*-tests (two tailed) (see [Figure S7A](#) for the *P*-maps for this analysis). Cluster-level statistics and permutation distributions were computed by permuting the conditions' labels in the original dataset 1000 times. The sample size for these analyses was 36 participants.

Nesting analysis

To assess the cross-frequency coupling between the phase of low-frequency neural oscillations, and the amplitude of high-frequency oscillations, we subjected each non-rejected epoch to a continuous wavelet transform using the MATLAB toolbox 'WTools'.²⁶ This process is similar to the one described in the [time-frequency analysis](#), however, this time we obtained the time-frequency maps for the amplitude information as well as for the phase information at the epoch level, at 1 Hz intervals in the 1–50 Hz range. To remove the distortion introduced by the wavelet transform, we removed 200 ms at each edge of the epoch's time-frequency responses for phase and amplitude. Then we performed baseline correction for the amplitude information at the epoch level. This process yielded two time-frequency maps for each epoch, channel, and condition, at the participant level: a map of spectral amplitude values, and a map of phase information.

Modulation index calculation

Using the phase and amplitude information obtained at the epoch level, we computed the modulation index M as a measure of coupling between the two signals. To do so, we followed the procedure described by Canolty and colleagues,⁴⁹ which has been found to be robust to various modulating factors and to successfully reconstruct coupling when tested on simulated EEG data as the ground truth.⁶⁸ To determine the modulation index we constructed a composite complex signal by combining the amplitude information of one frequency with the phase information of another frequency. Each sample of this composite signal represents a point in the complex plane. Measuring the degree of asymmetry of the probability density function (PDF) of the composite signal, as can be done by computing the mean of the composite signal, provides a measure of coupling between its amplitude and its phase. We constructed composite complex signals with the phase information of the lower-frequencies (1–30 Hz) and the amplitude information of the higher-frequencies (1–50 Hz) to obtain the modulation index as a function of analytic phase and analytic amplitude. At each frequency, we normalized the amplitude information to values between 0 and 1 before computing the modulation index. For the modulation index to be used as a metric of coupling strength, it must first be normalized. To accomplish this, we compared the mean M (M_{RAW}) to a set of permuted means (M_{PERM}) created by permuting the amplitude information 100 times. The modulus or length of M_{RAW} , compared to the distribution of permuted lengths, provides a measure of the coupling strength. The modulation index measure as defined by Canolty and colleagues,⁴⁹ corresponds to the normalized length $M_{\text{NORM}} = (M_{\text{RAW}} - \mu) / \sigma$, where μ is the mean of the permuted lengths and σ their standard deviation. We computed this modulation index measure for each non-rejected epoch, channel and condition individually, and then we averaged it across non-rejected epochs.

To assess whether cross-frequency coupling differed across conditions, we submitted the modulation index values from the initial and final resting state periods, as well as from the three language conditions to repeated measures ANOVAs with the within-subject factor Condition (InitialRest/French/Spanish/English/FinalRest). For this analysis each channel, and time/frequency sample was tested individually for the frequency ranges identified as relevant (1–8 Hz frequency for phase and 1–30 Hz frequency for amplitude) (see [Figure S9](#) for the P -maps for this analysis). To correct for multiple comparisons, we performed a permutation analysis, as the one described in previous sections. The sample size for these analyses was 36 participants.