



## OPEN ACCESS

## EDITED BY

Judit Gervain,  
Centre National de la Recherche Scientifique  
(CNRS), France

## REVIEWED BY

Irene De La Cruz Pavia,  
Université Paris Cité, France  
Gaia Lucarini,  
University of Padua, Italy

## \*CORRESPONDENCE

Carles Escera

✉ cescera@ub.edu

Jordi Costa-Faidella

✉ jcostafaidella@ub.edu

<sup>†</sup>These authors have contributed equally to this work

RECEIVED 31 January 2024

ACCEPTED 22 April 2024

PUBLISHED 22 May 2024

## CITATION

Gorina-Careta N, Arenillas-Alcón S,  
Puertollano M, Mondéjar-Segovia A,  
Ijjou-Kadiri S, Costa-Faidella J,  
Gómez-Roig MD and Escera C (2024)

Exposure to bilingual or monolingual  
maternal speech during pregnancy affects the  
neurophysiological encoding of speech  
sounds in neonates differently.

*Front. Hum. Neurosci.* 18:1379660.

doi: 10.3389/fnhum.2024.1379660

## COPYRIGHT

© 2024 Gorina-Careta, Arenillas-Alcón,  
Puertollano, Mondéjar-Segovia, Ijjou-Kadiri,  
Costa-Faidella, Gómez-Roig and Escera. This  
is an open-access article distributed under  
the terms of the [Creative Commons  
Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use,  
distribution or reproduction in other forums is  
permitted, provided the original author(s) and  
the copyright owner(s) are credited and that  
the original publication in this journal is cited,  
in accordance with accepted academic  
practice. No use, distribution or reproduction  
is permitted which does not comply with  
these terms.

# Exposure to bilingual or monolingual maternal speech during pregnancy affects the neurophysiological encoding of speech sounds in neonates differently

Natàlia Gorina-Careta<sup>1,2,3†</sup>, Sonia Arenillas-Alcón<sup>1,2,3†</sup>,  
Marta Puertollano<sup>1,2,3</sup>, Alejandro Mondéjar-Segovia<sup>1,2</sup>,  
Siham Ijjou-Kadiri<sup>1,2</sup>, Jordi Costa-Faidella<sup>1,2,3\*</sup>,  
María Dolores Gómez-Roig<sup>3,4</sup> and Carles Escera<sup>1,2,3\*</sup>

<sup>1</sup>Brainlab – Cognitive Neuroscience Research Group, Departament de Psicologia Clínica i Psicobiologia, Universitat de Barcelona, Barcelona, Spain, <sup>2</sup>Institut de Neurociències, Universitat de Barcelona, Barcelona, Spain, <sup>3</sup>Institut de Recerca Sant Joan de Déu, Esplugues de Llobregat, Barcelona, Spain, <sup>4</sup>BCNatal – Barcelona Center for Maternal Fetal and Neonatal Medicine (Hospital Sant Joan de Déu and Hospital Clínic), University of Barcelona, Barcelona, Spain

**Introduction:** Exposure to maternal speech during the prenatal period shapes speech perception and linguistic preferences, allowing neonates to recognize stories heard frequently *in utero* and demonstrating an enhanced preference for their mother's voice and native language. Yet, with a high prevalence of bilingualism worldwide, it remains an open question whether monolingual or bilingual maternal speech during pregnancy influence differently the fetus' neural mechanisms underlying speech sound encoding.

**Methods:** In the present study, the frequency-following response (FFR), an auditory evoked potential that reflects the complex spectrotemporal dynamics of speech sounds, was recorded to a two-vowel /oa/ stimulus in a sample of 129 healthy term neonates within 1 to 3 days after birth. Newborns were divided into two groups according to maternal language usage during the last trimester of gestation (monolingual; bilingual). Spectral amplitudes and spectral signal-to-noise ratios (SNR) at the stimulus fundamental (F<sub>0</sub>) and first formant (F<sub>1</sub>) frequencies of each vowel were, respectively, taken as measures of pitch and formant structure neural encoding.

**Results:** Our results reveal that while spectral amplitudes at F<sub>0</sub> did not differ between groups, neonates from bilingual mothers exhibited a lower spectral SNR. Additionally, monolingually exposed neonates exhibited a higher spectral amplitude and SNR at F<sub>1</sub> frequencies.

**Discussion:** We interpret our results under the consideration that bilingual maternal speech, as compared to monolingual, is characterized by a greater complexity in the speech sound signal, rendering newborns from bilingual mothers more sensitive to a wider range of speech frequencies without generating a particularly strong response at any of them. Our results contribute to an expanding body of research indicating the influence of prenatal experiences on language acquisition and underscore the necessity of including prenatal language exposure in developmental studies on language acquisition, a variable often overlooked yet capable of influencing research outcomes.

## KEYWORDS

speech brainstem responses, bilingualism, newborns, early language acquisition, frequency-following response (FFR), prenatal exposure

## Introduction

The process of language acquisition has long been a point of uncertainty in research exploring the roots of human language. Researchers have conducted extensive investigations to understand the initial state and process of language acquisition, providing insights into how environmental and genetic factors interact to fashion language and cognitive function, and the mechanisms underlying brain plasticity (Weaver et al., 2004; Werker and Tees, 2005; Barkat et al., 2011; Werker and Hensch, 2015). It is now widely accepted that both genetic and experiential factors contribute to language acquisition (Werker and Curtin, 2005; Gervain and Mehler, 2010), and researchers are interested in understanding how these factors interact during human development.

Infants at birth already exhibit advanced speech perception and language learning abilities. Newborns manifest a preference for speech over non-speech sounds (Vouloumanos and Werker, 2007), can discriminate between different languages based on their speech rhythms (Ramus et al., 2000), detect word boundaries (Christophe et al., 2001), discriminate words with different patterns of stress (Sansavini et al., 1997), or even distinguish consonant sounds (Cabrera and Gervain, 2020) and encode voice pitch in an adult-like manner (Arenillas-Alcón et al., 2021). These findings support the role of a genetically driven cerebral organization towards processing specific speech characteristics.

However, the prenatal period is not devoid of language experience and the study of its influence on the newborn's speech and language encoding capacities is receiving increasing attention. Hearing becomes functional and undergoes most of its development around the 26th to 28th week of gestation, allowing the fetus to perceive the maternal speech signal (Ruben, 1995; Moore and Linthicum, 2007; Granier-Deferre et al., 2011; May et al., 2011; Anbuhl et al., 2016). Although the exact characteristics of the acoustic signal reaching the fetus are not fully understood, intrauterine recordings from animal models and simulations suggest that the maternal womb acts as a low-pass filter, attenuating around 30 dB for frequencies over 600–1,000 Hz (Gerhardt and Abrams, 2000). The low-frequency components of speech that are transmitted through the uterus include pitch, slow aspects of rhythm and some phonetic information (Moon and Fifer, 2000; May et al., 2011). Evidence indicates that prenatal exposure to speech, despite attenuated by the filtering properties of the womb, shapes speech perception and linguistic preferences of newborns, as shown by studies revealing that neonates can recognize a story heard frequently *in utero* (DeCasper and Spence, 1986), prefer the voice of their mother (DeCasper and Fifer, 1980) and prefer their native language (Moon et al., 1993). Additionally, prenatal learning extends beyond these common preferences. Recent findings indicate that infants acquire specific knowledge of the prosody

(Gervain, 2018) and prefer the rhythmic patterns of the language they were exposed to while *in utero* (Mariani et al., 2023), indicating a very early specialization for their native language.

Yet, with reported rates of bilingualism of around 65% in Europe (Luk, 2017), an open question remains on the influence of prenatal exposure to more than one language on neural plasticity. Over the past 20 years, mounting evidence has suggested that both exposure to a bilingual acoustic environment and learning several languages affects not only language acquisition but a wide range of developmental processes including perception, cognition and brain development (Byers-Heinlein et al., 2019). Prior research has highlighted that early exposure to language influences infants' acquisition of speech sounds, indicating that, at birth, infants are able to discriminate all phonetic contrasts. As infants age, their perceptual systems are tuned to collapse over phonetic contrasts not found in the input language or languages, such that their ability to distinguish between phonetic elements becomes increasingly specific to their native language(s) (Kuhl et al., 2006; Saffran et al., 2006; Gervain and Werker, 2008; Kovács and Mehler, 2009; Bosch and Sebastián-Gallés, 2010). Moreover, cross-language interactions modulate almost every level of language processing, including speech perception, phonological, vocabulary and semantic development [for comprehensive review, refer to Hammer et al. (2014) and Kroll et al. (2012)]. Furthermore, some bilinguals switch from one language to the other within the same sentence, demonstrating greater demands on cognitive control than monolinguals to navigate the potential cross-language competition considering that language production is equivalent (Kovács and Mehler, 2009).

Speaking two languages daily also has consequences for the way in which higher cognitive processes operate and results in more precocious development of inhibition and attentional abilities (Costa et al., 2008; Kovács and Mehler, 2009; for review see Barac et al., 2014; Bialystok, 2017). There is evidence for functional and structural brain changes associated with bilingualism, even after brief periods of second-language learning (for extensive review see Li et al., 2014). Bilingual infants show different brain responses to native and non-native speech sounds than monolingual infants (Conboy and Kuhl, 2011). Bilingualism also affects the structure of both grey (Ressel et al., 2012) and white matter (Kuhl et al., 2016) in adults. The observed advantages in cognitive control and attentional abilities, as well as the pattern of structural differences, are modulated by the age of second language acquisition, whether the two languages were acquired simultaneously from birth or sequentially later in life and the interaction between languages (Kroll et al., 2012; Barac et al., 2014; Li et al., 2014).

As bilingual mothers speak using two different sets of phonemic categories and even use two slightly different voice pitch ranges (e.g., Ordin and Mennen, 2017), *in-utero* bilingual environments are characterized by a greater complexity of the reaching speech signal than monolingual ones. Interestingly, neonates exposed prenatally to

a bilingual environment can discriminate their two native languages already at birth and exhibit equal preferences for both (Byers-Heinlein et al., 2010). Thus, it appears clear that linguistic experiences while *in utero* play a significant role in shaping the early development of speech processing. However, how different prenatal maternal linguistic exposure influences the neural mechanisms underlying speech sound processing at birth is currently unknown.

A large body of evidence has supported the study of the neural encoding of speech sounds through electrophysiological recordings. In particular, the frequency-following response (FFR) can provide insights into the underlying neural mechanisms associated with prenatal language experience, shedding light on how early linguistic exposure shapes the speech-encoding capacities of newborns. The FFR is an auditory evoked potential elicited by periodic complex sounds that reflects neural synchronization with the auditory eliciting signal along the ascending auditory pathway (Skoe and Kraus, 2010; Krizman and Kraus, 2019), providing an accurate snapshot of the neural encoding of speech sounds. FFR recordings have thus become a useful tool to investigate the ability to distinguish between the pitch of different speakers' voices and the ability to encode the fine spectrotemporal details that distinguish different voiced speech sounds (Gorina-Careta et al., 2022). The interest in the neonatal FFR arises from its potential to serve as a predictive measure for future language development (Schochat et al., 2017), since alterations in FFR patterns in children have been associated with difficulties in reading and learning, dyslexia, impairments in phonological awareness and even autism (King et al., 2002; Banai et al., 2009; Chandrasekaran et al., 2009; Basu et al., 2010; Hornickel et al., 2012; Lam et al., 2017; Otto-Meyer et al., 2018; Font-Alaminos et al., 2020; Rosenthal, 2020). Interestingly, the FFR reflects the impact of a wide range of auditory experiences in children and adults, including training interventions, musical practice and bilingualism (Russo et al., 2005; Song et al., 2008; Kraus and Chandrasekaran, 2010; Krizman et al., 2012, 2015; Carcagno and Plack, 2017; Skoe et al., 2017; Gorina-Careta et al., 2019). In adults it has been observed that bilingual experience enhances the neural responses to the fundamental frequency of sounds (Krizman et al., 2015; Skoe et al., 2017), as well as the subcortical representation of pitch-relevant information (Krizman et al., 2012) and neural consistency, which correlated with both a better attentional control and language proficiency (Krizman et al., 2014). In neonates, FFR recordings have also been used to study the effects at birth of prenatal fetal auditory experiences such as music exposure (Arenillas-Alcón et al., 2023), but the influence of prenatal maternal bilingual speech remains unexplored.

In the present study, we aimed to examine the influence of maternal bilingual linguistic exposure *in-utero* in speech sound encoding at birth. To that end, we recorded FFRs from newborns who had been exposed to either a monolingual or a bilingual fetal environment during the last trimester of gestation and analyzed their capacity to encode voice pitch and vocalic formant structure information.

## Methods

### Participants

A sample of 131 newborns (mean age after birth =  $38.32 \pm 23.8$  h) was recruited from *SJD Barcelona Children's Hospital* in Barcelona

(Spain) and divided into two groups based on a short retrospective questionnaire delivered to the babies' mothers. Mothers were asked if they communicated using more than one language during the last 3 months of pregnancy and were instructed to report which languages they communicated in, provided they accounted for a minimum of 20% language usage time. Based on the collected responses, a total of 53 newborns were assigned to the group exposed to a monolingual fetal acoustic environment (MON; 27 females; mean gestational age =  $39.93 \pm 1.03$  weeks; mean birth weight =  $3,321 \pm 272$  g). A total of 76 newborns were assigned to the bilingual-exposed group (BIL; 33 females; mean gestational age =  $39.71 \pm 0.99$  weeks; mean birth weight =  $3,328 \pm 327$  g) after excluding two newborns, as their mothers were multilingual in Spanish, Catalan and English, being the third language used  $\geq 20\%$  of the time. Regarding the languages spoken by the bilingual mothers, all except one were Spanish—Other language and most of them were Spanish-Catalan bilinguals (77.3%). The other languages spoken were Arabic (6/75), English (1/75), Galician (1/75), German (1/75), Italian (2/75), Portuguese (2/75), Guaraní (2/75) and Romanian (2/75). On the other hand, newborns in the monolingual group were either exposed to Spanish (90.6%) or Catalan (9.4%).

No significant differences were found across groups in gestational age ( $U_{(127)} = 1868.500$ ,  $p = 0.370$ ), birth weight ( $t_{(127)} = -0.116$ ,  $p = 0.908$ ) and sex ( $\chi^2 = 0.710$ ,  $p = 0.399$ ). Maternal education level and musical exposure were assessed using a sociodemographic questionnaire (an English version of the sociodemographic questionnaire can be found in the [Supplementary material](#)). Groups did not differ in maternal educational level ( $\chi^2 = 1.992$ ,  $p = 0.574$ ), a key confounding factor associated with language acquisition and development (Hoff, 2003; Rowe, 2008) closely tied to the linguistic environment a fetus is exposed to. We also ascertained that groups did not differ in prenatal musical exposure [ $\chi^2 = 0.025$ ,  $p = 0.874$ ; see Arenillas-Alcón et al. (2023) for details], as it exerts a significant impact on speech encoding capacities at birth (Partanen et al., 2013b, 2022; Arenillas-Alcón et al., 2023).

All neonates obtained Apgar scores higher than 8 at 1 and 5 min of life and passed adequately the universal newborn hearing screening (UNHS) before the recruitment. According to the recommendations of the [Joint Committee on Infant Hearing \(2019\)](#), newborns born from high-risk gestations, after obstetric pathologies or any other kind of risk factors related to hearing impairment were excluded from the recruitment.

Additionally, as performed in previous research from our laboratory (Ribas-Prats et al., 2019, 2021, 2023; Arenillas-Alcón et al., 2021, 2023), both groups of newborns received a standard click-evoked auditory brainstem response (ABR) test to ensure the integrity of the auditory pathway. A click-stimulus, with a duration of 100  $\mu$ s, was employed during the test, presented at a rate of 19.30 Hz with an intensity of 60 dB sound pressure level (SPL) until a total of 4000 artifact-free repetitions were collected. A prerequisite for participation in the experiment for all newborns was the successful identification of the wave V peak. This study was approved by the Ethical Committee of Clinical Research of the Sant Joan de Déu Foundation (Approval ID: PIC-53-17), and required the mothers to fill out a sociodemographic questionnaire and to sign an informed consent prior to the participation, in line with the Code of Ethics of the World Medical Association (Declaration of Helsinki).

## Stimulus

Neonatal FFRs were collected to a two-vowel stimulus with a rising pitch ending (/oa/; Arenillas-Alcón et al., 2021). The /oa/ stimulus was created in Praat (Boersma and Weenink, 2020) and had a total length of 250 ms divided into three different sections, according to its fundamental frequency ( $F_0$ ) and its formant content (/o/ vowel section: 0–80 ms,  $F_0 = 113$  Hz,  $F_1 = 452$  Hz,  $F_2 = 791$  Hz; /a/ formant transition section = 80–90 ms; /a/ vowel steady section = 90–160 ms,  $F_0 = 113$  Hz,  $F_1 = 678$  Hz,  $F_2 = 1,017$  Hz; /a/ vowel rising section = 160–250 ms,  $F_0 = 113$ –154 Hz,  $F_1 = 678$  Hz,  $F_2 = 1,017$  Hz; Figure 1A).

The stimulus was designed with optimal parameters to study the frequency-following response, specially taking into account that due to the low-pass filter characteristics of the womb, fetuses are isolated from the mid and high frequency acoustic content of external sounds that characterizes most of the temporal fine structure of speech. The /oa/ stimulus used includes a pitch variation and two vowel sections with different formant structure based on relatively lower frequency harmonic components and suitable durations for accurate spectral analyses, which enable a proper assessment of speech sound temporal envelope and temporal fine structure encoding (Krizman and Kraus, 2019; Arenillas-Alcón et al., 2021). The relatively low  $F_0$  frequency, typical of a male speaker, was chosen to ensure a reliable measure of the neural representation of sound pitch (Krizman and Kraus, 2019) and the phonetic contrasts (/o/; /a/) belong to the phonetic repertoire of both Spanish and Catalan languages.

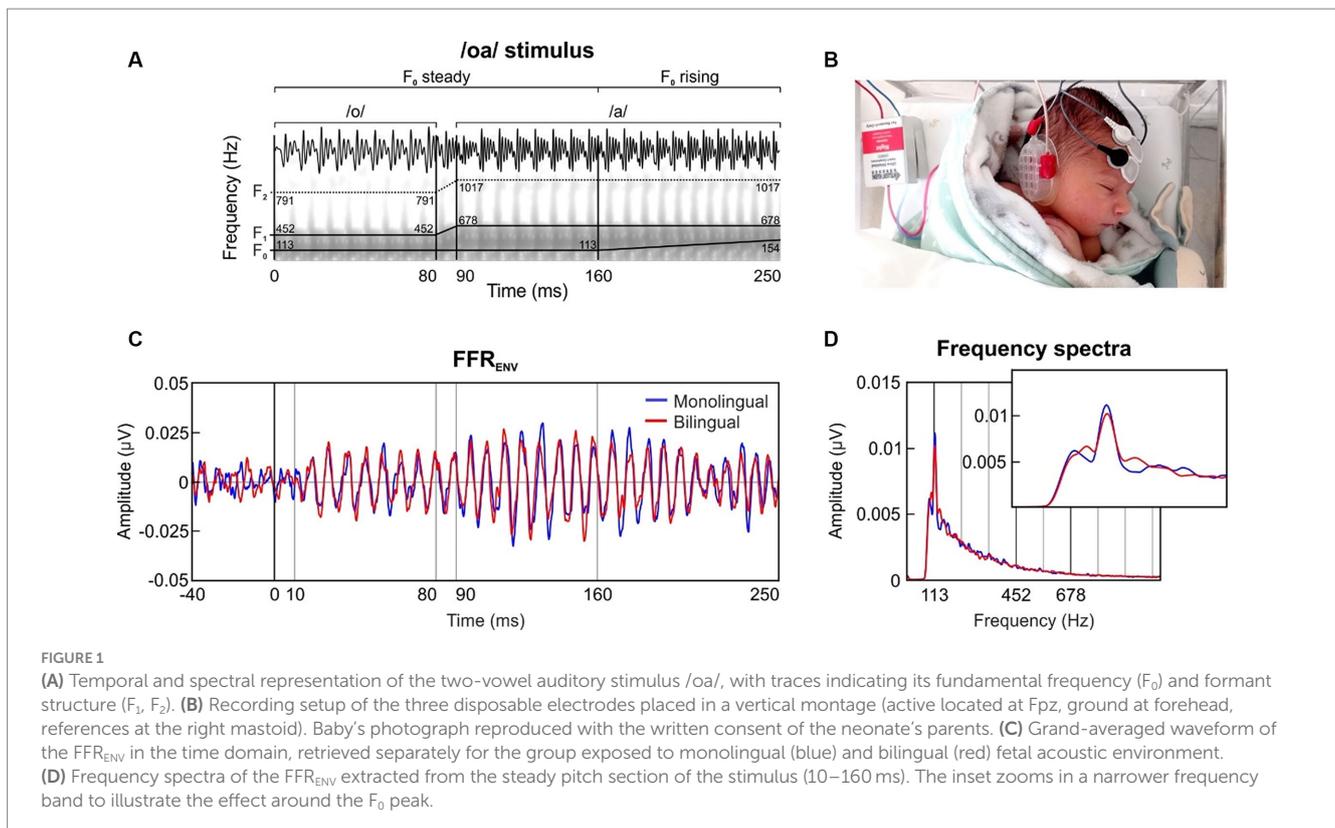
The /oa/ stimulus was presented at a rate of 3.39 Hz in alternating polarities and delivered monaurally to the right ear at 60 dB SPL of

intensity with an earphone connected to a Flexicoupler disposable adaptor (Natus Medical Incorporated, San Carlos, CA).

## Procedure and data acquisition

After the successful completion of the UNHS, neonates were tested at the hospital room while they were sleeping in their bassinet. Three disposable Ag/AgCl electrodes were placed in a vertical montage configuration (active at Fpz, ground at forehead, reference at the right mastoid, ipsilateral to the auditory stimulation; as shown in Figure 1B), ensuring impedances below 7 k $\Omega$ . The presentation of click and speech stimuli was done by using a SmartEP platform connected to a Duet amplifier, which incorporated the cABR and the Advanced Hearing Research modules (Intelligent Hearing Systems, Miami, FL, United States).

The experimental procedure involved the recording of two blocks of click stimuli, followed by four blocks of 1000 artifact-free responses to the /oa/ stimulus. Any electrical activity surpassing  $\pm 30$   $\mu$ V threshold was automatically rejected until a total of 4,000 presentations was collected. The total mean duration of the recording session was approximately 25 min [2 click blocks  $\times$  2,000 repetitions  $\times$  51.81 ms SOA + 4 /oa/ blocks  $\times$  1,000 repetitions  $\times$  295 ms of stimulus-onset asynchrony (SOA)] including the duration of rejected sweeps. The continuous EEG signal was acquired at a sampling rate of 13,333 Hz with an online bandpass filter with cutoff frequencies from 30 to 1,500 Hz and online epoched from  $-40.95$  ms (pre-stimulus period) to 249.975 ms.



## Data processing and analysis

Data epochs were bandpass filtered offline from 80 to 1,500 Hz and averaged separately per stimulus polarity. To highlight the encoding of the stimulus fundamental frequency ( $F_0$ ) and to reduce the contribution of cochlear microphonics, neural responses to the two opposite stimulus polarities were added  $[(\text{Condensation} + \text{Rarefaction})/2]$ , obtaining the envelope-following response ( $\text{FFR}_{\text{ENV}}$ ). Further, to emphasize the FFR components associated with the encoding of the stimulus temporal fine structure, such as the first formant ( $F_1$ ), while reducing the impact of envelope-related activity, the neural responses to alternating polarities were subtracted  $[(\text{Condensation} - \text{Rarefaction})/2]$ , yielding the temporal fine structure-following response ( $\text{FFR}_{\text{TFS}}$ ; Aiken and Picton, 2008; Krizman and Kraus, 2019). Considering the stimulus formant content, we focused our analyses exclusively on the spectral peaks that corresponded to  $F_1$  frequencies, as  $F_2$  frequencies fall at the limits of the spectral resolution of the FFR, resulting in elicited neural responses relatively weak and challenging to be accurately observed in newborns (Gorina-Careta et al., 2022). Detailed information regarding the analyzed parameters from the neonatal FFR can be found below. All parameters were computed using custom scripts in Matlab R2019b (The Mathworks Inc., 2019), developed in our laboratory and previously employed in similar analyses in former studies (Arenillas-Alcón et al., 2021).

### Neural lag

Neural lag served as an indicator of the neural transmission delay within the auditory system, and was assessed to estimate the time passed from cochlear stimulus reception to the onset of neural phase-locking (Jeng et al., 2010; Liu et al., 2015; Ribas-Prats et al., 2019, 2021, 2023; Arenillas-Alcón et al., 2021, 2023). To calculate the neural lag, a cross-correlation analysis was computed between the auditory stimulus and the neural response. The neural lag was determined by identifying the time lag corresponding to the highest cross-correlation value within a time window of 3–13 ms.

### Pre-stimulus root mean square (RMS) amplitude

The RMS of the pre-stimulus period was employed as a measure of the general magnitude of neural activity over time, and to dismiss electrophysiological disparities in the pre-stimulus region (Liu et al., 2015; White-Schwoch et al., 2015; Ribas-Prats et al., 2019, 2021, 2023; Arenillas-Alcón et al., 2023). This measure was computed by squaring each data point within the pre-stimulus region of the neural response (from  $-40$  to  $0$  ms), calculating the mean of the squared values and subsequently obtaining the square root of the resulting average.

## Voice pitch encoding from $\text{FFR}_{\text{ENV}}$

### Spectral amplitude at $F_0$

Spectral amplitude at  $F_0$  (113 Hz) was used as a quantitative measure of the neural phase-locking strength at the specific frequency of interest (White-Schwoch et al., 2015; Ribas-Prats et al., 2019, 2021, 2023; Arenillas-Alcón et al., 2021, 2023). It was computed by applying a fast Fourier transform (FFT; Cooley and Tukey, 1965) to obtain the frequency spectrum of the neural response during the steady pitch section of the stimulus (10–160 ms), and then calculating the average

amplitude within a  $\pm 5$  Hz window centered around the peak of the stimulus  $F_0$ .

### Signal-to-noise ratio at $F_0$

Signal-to-noise ratio (SNR) at  $F_0$  was analyzed to obtain an estimation of the relative spectral magnitude of the response, taking into account not only to the amplitude value at the  $F_0$  frequency peak (113 Hz) but also the noise levels at the surrounding frequencies. Therefore, the SNR was calculated by dividing the mean spectral amplitude within a  $\pm 5$  Hz frequency window centered at the peak of the frequency of interest (113 Hz) by the averaged mean amplitude within two additional 28 Hz wide frequency windows (flanks), centered at  $\pm 19$  Hz from the frequency of interest (80–108 Hz and 118–146 Hz).

## Formant structure encoding from $\text{FFR}_{\text{TFS}}$

### Spectral amplitudes at $F_1$ peaks

To assess spectral amplitudes at the specific spectral peaks regarding the stimulus  $F_1$  frequencies (452 Hz [ $/o/$ ] and 678 Hz [ $/a/$ ]), the neural responses corresponding to the  $/o/$  section (10–80 ms time window) and the  $/a/$  steady section (90–160 ms time window) were individually analyzed and the respective amplitudes within a  $\pm 5$  Hz window centered at the peak frequencies corresponding to the vowel formant centers were extracted. The transition from  $/o/$  vowel to  $/a/$  vowel was not analyzed due to its short duration (10 ms).

### Signal-to-noise ratio at $F_1$

To compute the relative spectral magnitude of the response at the stimulus  $F_1$  frequencies considering noise levels, SNRs at spectral peaks that correspond to the stimulus  $F_1$  frequencies (452 Hz and 678 Hz) were calculated separately on the  $/o/$  and the  $/a/$ -steady sections. To do so, the SNR was calculated by dividing the mean spectral amplitude within a  $\pm 5$  Hz frequency window centered at the peak of the frequency of interest (452 or 678 Hz) by the averaged mean amplitude within two additional 28 Hz wide frequency windows (flanks), centered at  $\pm 26$  Hz from the frequency of interest (for 452 Hz peak: 402–430 Hz and 474–502 Hz; for 678 Hz peak: 628–656 Hz and 700–728 Hz).

## Statistical analysis

Statistical analyses were conducted using Jamovi 2.3.26 (The Jamovi Project, 2023). Descriptive statistics were calculated, including the mean, standard deviation (SD), median, first quartile ( $Q_1$ ), third quartile ( $Q_3$ ), interquartile range (IQR), and minimum and maximum values, for each computed parameter within the two groups of newborns (MON; BIL).

To analyze the effects of prenatal bilingual exposure on neural transmission delay, pre-stimulus root mean square amplitude and voice pitch encoding depending on the normality of the data, two-tailed independent samples  $t$ -tests or Mann-Whitney U tests were conducted to evaluate significant differences between groups, with Cohen's  $d$  being reported as the effect size. Kolmogorov-Smirnov test was used to assess the normal distribution of the samples.

The effects of prenatal bilingual exposure on formant structure encoding were analyzed with two repeated-measures ANOVAs with the factor Stimulus Section (/o/ section; /a/ section) as within-subjects factor and the factor Group (Monolingual; Bilingual) as between-subjects factor for each of the two formant amplitudes (452 and 678 Hz) separately. The Greenhouse–Geisser correction was applied when the assumption of sphericity was violated. Additional two-tailed independent samples Mann–Whitney U post-hoc tests were performed to examine the direction of the effects. Results were considered statistically significant when  $p < 0.05$ .

## Results

Frequency following responses (FFR) elicited by a two-vowel speech stimulus /oa/ (Figure 1A) were collected from a total sample of 129 newborns divided into two groups according to their prenatal fetal exposure to monolingual (MON) or bilingual (BIL) maternal speech. To comprehensively evaluate the neonates' ability to encode the pitch and vowel formant structure of speech sounds, the neural responses to the fundamental frequency ( $F_0$ ) and the vowels' first formant ( $F_1$ ) were analyzed considering the distinct sound characteristics of the different stimulus sections. All detailed descriptive statistics from the parameters analyzed can be found in Supplementary Table S1.

### Neural transmission delay

No significant differences were found across groups in neural lag ( $U_{(127)} = 1950.500$ ,  $p = 0.763$ , Rank-biserial correlation = 0.032).

### Pre-stimulus root mean square (RMS) amplitude

There were no statistically significant differences observed between the groups with regards to the background neural activity preceding the auditory stimulation ( $U_{(127)} = 1914.000$ ,  $p = 0.634$ , Rank-biserial correlation = 0.050).

### Voice pitch encoding ( $FFR_{ENV}$ )

The grand-averaged  $FFR_{ENV}$  waveform for each group is illustrated in Figure 1C. To assess the robustness of the voice pitch representation, we analyzed the steady section (10–160 ms) of the /oa/ stimulus with a steady fundamental frequency ( $F_0$ ) of 113 Hz.

The grand-averaged spectral representation of the neonatal FFR extracted from each group is depicted in Figure 1D. No differences were found across groups in spectral amplitude at  $F_0$  computed using the steady pitch section of the stimulus ( $U_{(127)} = 1736.000$ ,  $p = 0.184$ , Rank-biserial correlation = 0.138).

Yet, the statistical analyses performed on the  $F_0$  SNR, which represents the  $F_0$  relative spectral amplitude in relation with the spectral amplitude of the neighboring frequencies, revealed significant differences between groups, indicating that newborns exposed to a monolingual prenatal fetal environment exhibited significantly larger

SNR values as compared to the bilingual exposed neonates ( $U_{(127)} = 1508.000$ ,  $p = 0.016$ , Rank-biserial correlation = 0.251).

### Formant structure encoding ( $FFR_{TF5}$ )

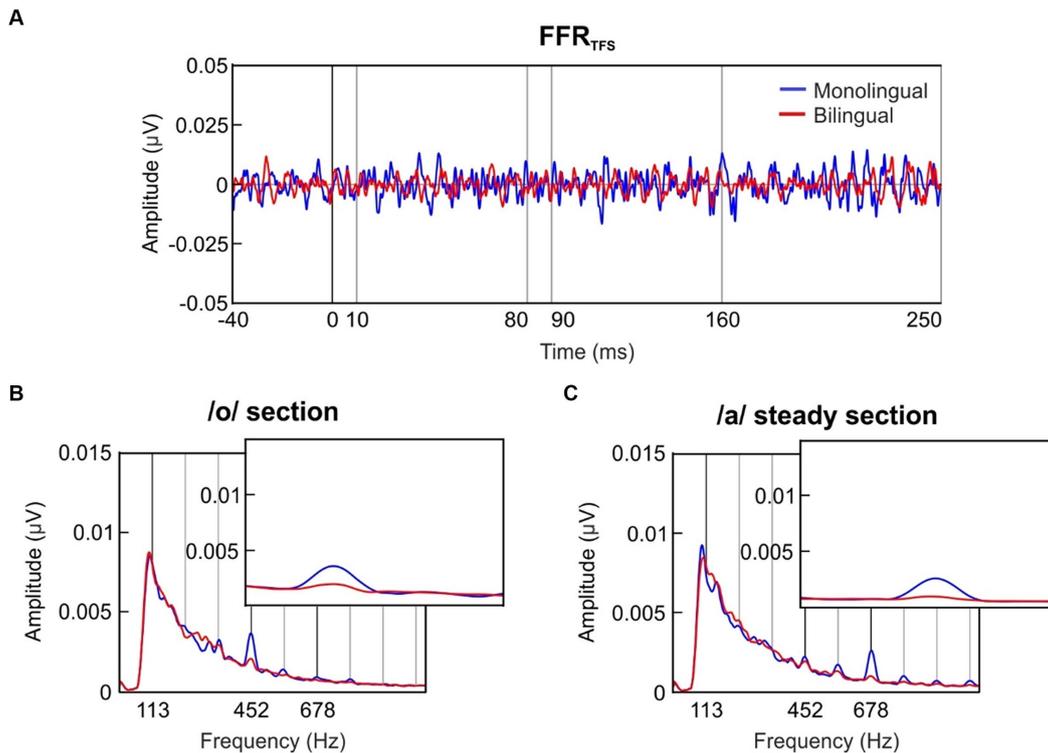
The grand-averaged  $FFR_{TF5}$  waveform for each group is shown in Figure 2A. To evaluate the newborns' ability to encode the formant structure of speech sounds, the /oa/ stimulus included two sections with the same voice pitch but different fine-structure. Specifically, the /o/ section (10–80 ms) was characterized by a center formant frequency ( $F_1$ ) of 452 Hz, and the /a/ steady section (90–160 ms) by a  $F_1$  frequency of 678 Hz. Spectral amplitudes were retrieved from the  $FFR_{TF5}$  separately from neural responses during the /o/ section and the /a/ steady-pitch section, selecting the spectral peaks corresponding to stimulus  $F_1$  frequencies.

The grand-averages of the  $FFR_{TF5}$  spectral amplitudes during the /o/ section are illustrated in Figure 2B for each group separately, while the spectral representations during the /a/ steady section are depicted in Figure 2C.  $F_1$  spectral amplitudes during the /o/ section and the /a/ steady section are depicted in Figure 3 for each group at each formant center frequency (452 Hz, 678 Hz) separately.

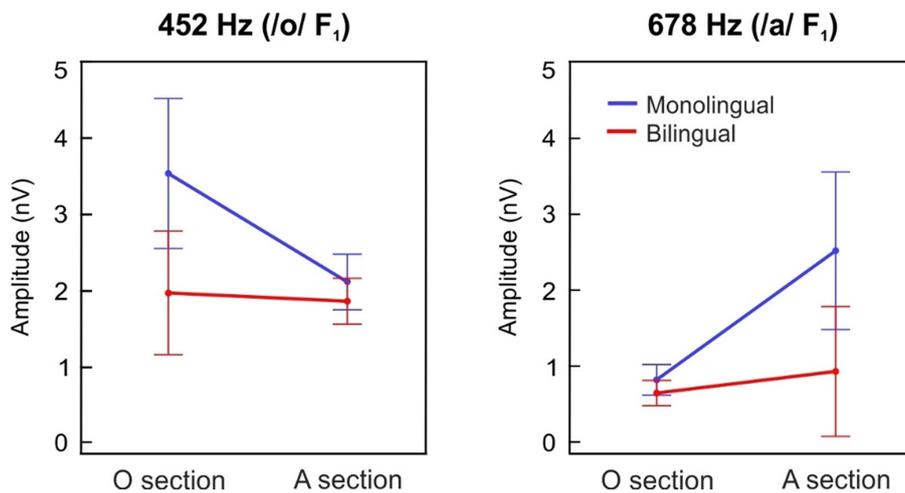
When analyzing the effects of a prenatal maternal bilingual language exposure in formant spectral amplitude at 452 Hz (Figure 3, left panel), which corresponds to the  $F_1$  center frequency of the /o/ vowel, a main effect of group revealed significantly greater spectral amplitudes in the MON group as compared to the BIL (group main effect;  $F_{(1,127)} = 4.939$ ,  $p = 0.028$ ,  $\eta^2 = 0.037$ ). Moreover, a significantly larger spectral amplitude was observed during the /o/ section vs. /a/ steady section (stimulus section main effect;  $F_{(1,127)} = 7.580$ ,  $p = 0.007$ ,  $\eta^2 = 0.056$ ), thus indicating a proper encoding of the vowel /o/ in its corresponding stimulus section. Interestingly, a significant interaction of group per stimulus section was identified as well (interaction;  $F_{(1,127)} = 5.809$ ,  $p = 0.017$ ,  $\eta^2 = 0.044$ ), demonstrating that MON neonates showed significantly larger spectral amplitudes during the /o/ section at its corresponding formant frequency than BIL.

Similar results were observed when analyzing the effects of a prenatal maternal bilingual language exposure in the formant encoding at 678 Hz (Figure 3, right panel), which corresponds to the  $F_1$  center frequency of the /a/ vowel. A main effect of group revealed significantly greater spectral amplitudes in the MON group as compared to the BIL (group main effect;  $F_{(1,127)} = 5.01$ ,  $p = 0.027$ ,  $\eta^2 = 0.038$ ). Moreover, a significantly larger spectral amplitude at 678 Hz during the /a/ steady section vs. /o/ section was observed (stimulus section main effect;  $F_{(1,127)} = 10.93$ ,  $p = 0.001$ ,  $\eta^2 = 0.079$ ), thus indicating a proper encoding of the /a/ vowel in its corresponding stimulus section. Interestingly, a significant interaction of group per stimulus section was also identified (interaction;  $F_{(1,127)} = 5.812$ ,  $p = 0.017$ ,  $\eta^2 = 0.044$ ), demonstrating that the MON group exhibited higher spectral amplitudes during the /a/ steady section at its corresponding frequency than the BIL.

The same pattern of results was obtained when comparing the relative spectral amplitude of the response at the stimulus  $F_1$  frequencies taking into account the neural response to the neighboring frequencies. When analyzing the effects of a fetal maternal bilingual language exposure in SNR at 452 Hz, which corresponds to the  $F_1$  of the /o/ vowel, a main effect of group revealed significantly greater spectral amplitudes in the MON group as compared to the BIL (group



**FIGURE 2** Formant structure encoding. (A) Grand-averaged waveform of the  $FFR_{TFS}$  in the time domain, retrieved separately for the group exposed to a monolingual fetal acoustic environment (blue) and the bilingual-exposed group (red). (B) Frequency spectra of the  $FFR_{TFS}$  extracted from the /o/ section of the stimulus (10–80 ms). The inset zooms in a narrower frequency band to illustrate the effect around the /o/  $F_1$  peak (452 Hz) during the /o/ section. (C) Frequency spectra of the  $FFR_{TFS}$  extracted from the /a/ steady section of the stimulus (90–160 ms). The inset zooms in a narrower frequency band to illustrate the effect around the /a/  $F_1$  peak (678 Hz) during the /a/ steady section.



**FIGURE 3** Spectral amplitudes at the first formant ( $F_1$ ).  $F_1$  spectral amplitudes at 452 Hz (left) and 678 Hz (right) during the /o/ section (10–80 ms) and the /a/ steady section (90–160 ms), plotted in blue and red lines for the monolingual and the bilingual-exposed newborns, respectively. Error bars represent 95% confidence intervals.

main effect;  $F_{(1,127)} = 8.301, p = 0.005, \eta^2 = 0.061$ ). Moreover, a significantly larger spectral amplitude was observed during the /o/ section vs. /a/ steady section (stimulus section main effect;  $F_{(1,127)} = 7.517, p = 0.007, \eta^2 = 0.056$ ). A significant interaction of

group per stimulus section was identified as well (interaction;  $F_{(1,127)} = 7.304, p = 0.008, \eta^2 = 0.054$ ).

Similar effects were observed when analyzing the effects of a prenatal bilingual environment in the formant SNR at 678 Hz, which

corresponds to the frequency of the /a/ vowel. A main effect of group revealed significantly greater spectral amplitudes in the MON group as compared to the BIL (group main effect;  $F_{(1,127)} = 7.127, p = 0.009, \eta^2 = 0.053$ ). Moreover, a significantly larger spectral amplitude at 678 Hz during the /a/ steady section vs. /o/ section was observed (stimulus section main effect;  $F_{(1,127)} = 22.072, p < 0.001, \eta^2 = 0.148$ ). Finally, a significant interaction of group per stimulus section was also identified (interaction;  $F_{(1,127)} = 10.330, p = 0.002, \eta^2 = 0.075$ ).

## Discussion

The present study investigated the impact of maternal bilingual speech during pregnancy on the neural encoding of speech pitch and vowel formant structure in neonates. A total sample of 129 healthy-term newborns was divided into two groups according to their monolingual or bilingual prenatal exposure during the last trimester of gestation, as reported by their mothers through a questionnaire. FFRs elicited to a two-vowel speech stimulus /oa/ (Arenillas-Alcón et al., 2021) were recorded to assess the neural responses to the stimulus' fundamental frequency ( $F_0 = 113$  Hz; related to voice pitch encoding) and the first formant of each vowel (/o/  $F_1 = 452$  Hz; /a/  $F_1 = 678$  Hz; related to vowel formant structure encoding). Our results revealed that the neural representation of pitch, as indexed by the spectral amplitude of the  $FFR_{ENV}$  at the stimulus  $F_0$ , did not differ between monolingual and bilingual exposure groups, but monolingually exposed neonates exhibited a higher signal-to-noise ratio (SNR) at the  $F_0$  spectral peak, suggesting the contribution of a higher spectral noise at neighboring frequencies in the bilingual group. Additionally, monolingually exposed neonates exhibited larger spectral amplitudes and SNRs of the  $FFR_{TFS}$  at the formant peak frequencies ( $F_i$ ) of the speech stimulus used, indicating a stronger encoding of vocalic structure. Furthermore, no significant group differences were observed in neural lag and pre-stimulus root mean square (RMS) amplitude, implying comparable neural transmission delays and absence of a distinct overall neural activity prior to the auditory stimulation. Together, these findings provide novel insights into the effects of prenatal language exposure on the neural encoding of speech sounds at birth.

Pitch is a crucial attribute in the perception of periodic speech sounds, as it conveys prosodic information, facilitates speaker recognition and speech segmentation, accelerates phoneme acquisition in tonal languages, helps with language comprehension in noisy environments and even contributes to the perception of the emotional state in a conversation (Musacchia et al., 2007; Benavides-Varela et al., 2012; Partanen et al., 2013a; Plack et al., 2014; Gervain, 2018; Cabrera and Gervain, 2020; Arenillas-Alcón et al., 2021; Ribas-Prats et al., 2021). The fact that neural mechanisms underlying voice pitch encoding are already mature at birth (Jeng et al., 2011; Ribas-Prats et al., 2019; Cabrera and Gervain, 2020; Arenillas-Alcón et al., 2021) suggests that pitch may play a crucial role in the very first stages of language acquisition (Jeng et al., 2016). Going a step further, pitch could provide a neural synchrony channel onto which separate neural representations of other speech features would anchor as parts of an ensemble that would, ultimately, give rise to a coherent percept (Eggermont, 2001).

Previous studies demonstrated that pitch and pitch contour discrimination drastically improve with training (e.g., Carcagno and

Plack, 2017). In this regard, growing up in a bilingual environment, which is characterized as more demanding, dynamic, phonologically rich and requiring heightened attention to all linguistic input, is related to a strengthened neural representation of pitch (Krizman et al., 2012, 2015). Different languages have distinct overall height pitch levels. For example, Catalan was observed to have a higher pitch compared to Spanish (Marquina Zazura, 2011); Polish was found to have a higher pitch compared to American English (Majewski et al., 1972); Mandarin, a higher pitch than English (Keating and Kuo, 2012); Japanese, a higher pitch than Dutch (Van Bezooijen, 1995); or Slavic languages, a higher pitch than Germanic ones (Andreeva et al., 2014). Further, speakers of two phonologically similar dialects exhibit differences in their height pitch levels (e.g., two different dialects of Mandarin; Deutsch et al., 2009).

Yet, pitch height is not the only element that contributes significantly to the distinctiveness of a particular language. The intonational patterns, which are the rising and falling patterns of pitch that convey meaning and contribute to the rhythm of speech, may differ between the different languages. When a speaker switches between languages they naturally adjust the specific contours, pitch ranges, and other prosodic features to conform to the norms of the target language, and many linguistic features such as intonation, may affect the mean fundamental frequency of speech (Järvinen et al., 2013). This adjustment helps maintaining communicative clarity and aligns with the phonetic characteristics of the language being spoken (Mary and Yegnanarayana, 2008; Passoni et al., 2022).

With continued exposure to these complex linguistic contexts, the auditory system gradually becomes finely tuned to process sound more efficiently (Krizman et al., 2012). Thus, individuals with years of exposure and interaction with bilingual environments develop enhanced flexibility and speech-encoding abilities. Most notably, previous studies have shown that bilingual individuals, particularly females, exhibit different pitch frequency ranges depending on the language they speak (Ordin and Mennen, 2017). As both pitch and the intonational patterns of the languages are different, and the prosodic elements of speech which include pitch contours, rhythm, and stress (Moon and Fifer, 2000) are acoustic features reliably transmitted through the womb (Gerhardt and Abrams, 2000; May et al., 2011), bilingual mothers provide their children with a higher pitch variability *in utero*.

Considering the reviewed literature, if the developing auditory system of a fetus, who underwent approximately 3 months of noninteractional exposure to degraded speech, responded to acoustic exposure as the mature one, we would expect newborns from bilingual mothers to exhibit a higher neural encoding of voice pitch. But our results showed otherwise. We found no differences across groups in  $FFR_{ENV}$  spectral amplitudes at  $F_0$ , which aligns with the idea that pitch processing mechanisms are already mature at birth. Yet, we observed a decreased SNR at the  $F_0$  in newborns who were prenatally exposed to a bilingual environment. We attempt to reconcile our seemingly contradicting results by hypothesizing that the higher spectral amplitudes found in bilingually exposed neonates at  $F_0$  neighboring frequencies reflect an increased sensitivity to a wider range of pitch frequencies without yet generating a particularly strong response at any of them.

This view aligns with research on perceptual phonetic development, especially when growing in bilingual environments. Previous studies demonstrated that experience with language shapes

infants' abilities to process speech sounds and, with age, the newborn's ability to differentiate phonetic distinctions becomes more language-specific (Kuhl et al., 2006; Saffran et al., 2006; Gervain and Werker, 2008; Bosch and Sebastián-Gallés, 2010). At birth all infants possess the ability to perceive all sound distinctions used in languages as they are sensitive to the basic rhythmic differences between languages (Nazzi et al., 1998; Byers-Heinlein et al., 2010). Around 3–4 months of age infants are sensitive to rhythmic differences between languages that go beyond their belonging to the three basic rhythmic classes (Bosch and Sebastián-Gallés, 2010; Molnar et al., 2014) and by the age of 6 months monolingual infants' ability to perceive speech becomes tailored to their native language. Infants exposed to two languages are also able to discriminate the sound contrasts of both their languages, but this occurs only at the end of their first year (Bosch and Sebastián-Gallés, 2003; Sundara et al., 2008; for review see Hammer et al., 2014).

Yet, the early prenatal impact of language goes beyond language discrimination. As reviewed in the introduction, newborns prefer their mother's voice over other female voices (DeCasper and Fifer, 1980), their communicative cries reflect the prosody of the language they heard *in utero* (Mampe et al., 2009) and can recognize stories heard during pregnancy (DeCasper and Spence, 1986). Moreover, previous studies also demonstrated that differences in prenatal language exposure modulate perceptual grouping biases at birth (Abboub et al., 2016) and suggest that hearing pitch contrasts before birth may influence pitch-based grouping preferences and may lead to a stable bias at birth. Thus, despite the discrimination (or no discrimination) of languages at birth, prenatal language exposure modulates the processing of speech sounds. Our findings align with the suggested hypothesis that being bilingual confers a greater perceptual flexibility (Abboub et al., 2016), as we observed in bilingually exposed newborns an increased sensitivity to a wider range of pitch frequencies.

Our results also reveal a modulation of the neural encoding of vowel formants ( $F_1$ ) depending on prenatal linguistic exposure. In particular, monolingual-exposed neonates exhibited higher spectral amplitudes at the corresponding formant frequencies of the stimulus' /o/ and steady-/a/ vowels. In a previous study, we found that while the neural encoding of pitch was adult-like at birth, formant encoding was still immature (Arenillas-Alcón et al., 2021). As vowel formant center frequencies are language specific and stable regardless of voice pitch variation, which also presents slight modulations in monolingual individuals during natural speaking, the auditory system of a monolingual-exposed fetus receives a more consistent phonetic repertoire than that of a bilingual-exposed. This would possibly lead to a more effective and accurate encoding of the specific language vowel sound characteristics at birth. Simply put, monolingual newborns seem to have an advantage in processing the specific sounds of their mother tongue, a finding previously attributed to postnatal linguistic exposure (Kuhl, 2010). Our findings thus highlight the greater variability of acoustic speech inputs to which the fetus of bilingual mothers would be exposed and therefore suggest the need for bilinguals to develop a different phonological representation for each of the languages (Sebastian-Gallés et al., 2006). Further investigation into the developmental trajectories of auditory processing in different populations of newborns, with different prenatal auditory experiences, and using language-specific phonetic contrasts (e.g., Catalan contrasts such as /e - ε/), which are especially difficult –when not impossible– to detect for

Spanish-monolinguals (Pallier et al., 1997, 2001), may shed more light on this issue.

Despite being confident about our results due to the abovementioned reasons, we are fully aware of a number of limitations of our study: language exposure was assessed by a short (approx. 5 min answer time), retrospective questionnaire provided at the time of delivery, with a spoken description of the content of the questionnaire. This poses, at least, two factors not adequately controlled. First, the actual frequency in which mothers spoke any of the two languages, as we rely only on their reports referring to the last trimester of pregnancy. Furthermore, although a minimum period of usage time had to occur to be considered as valid, the questionnaire did not address the exact amount of language usage within a day. Future studies should address these limitations, for instance, by collecting large amounts of data from a maternal diary of language usage during the last trimester of pregnancy and include an additional language abilities test (such as LEAP-Q; Marian et al., 2007) to evaluate the putative link between  $F_0$  encoding abilities in newborns and maternal language usage percentage.

Overall, our findings emphasize the potential importance of prenatal linguistic exposure in shaping the neural mechanisms underlying language acquisition and highlight the sensitivity of the FFR in capturing these subtle changes. The results add to a growing body of research that suggests a role for prenatal fetal experiences in modeling language acquisition (Moon et al., 2012; Partanen et al., 2013b; Gervain, 2015, 2018; Arenillas-Alcón et al., 2023). Furthermore, they also highlight the importance of considering prenatal language exposure in developmental studies about language acquisition, a factor that is not routinely measured and reported, and that may contribute to divergent findings.

## Conclusion

The present study contributes significant insights into the impact of prenatal bilingual exposure on the neural encoding of speech sounds at birth, thereby increasing our knowledge of the early stages of language acquisition. The observed differences in the encoding of voice pitch and formant structure depending on prenatal linguistic exposure highlight the remarkable plasticity and learning potential of the human brain even before birth, emphasizing the complex interaction between genetic and environmental factors in shaping our cognitive abilities and linguistic development.

## Data availability statement

The data supporting the conclusions of this article will be made available upon request by the authors, without undue reservation.

## Ethics statement

The studies involving humans were approved by Ethical Committee of Clinical Research of the Sant Joan de Déu Foundation (Approval ID: PIC-53-17). The studies were conducted in accordance with the local legislation and institutional requirements. Written

informed consent for participation in this study was provided by the participants' legal guardians/next of kin.

## Author contributions

NG-C: Writing – review & editing, Writing – original draft, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. SA-A: Writing – review & editing, Writing – original draft, Visualization, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. MP: Investigation, Writing – review & editing, Methodology. AM-S: Writing – review & editing, Methodology, Investigation. SI-K: Writing – review & editing, Methodology, Investigation. JC-F: Writing – review & editing, Supervision, Methodology, Conceptualization. MG-R: Writing – review & editing, Resources, Funding acquisition. CE: Writing – review & editing, Supervision, Resources, Methodology, Funding acquisition, Conceptualization.

## Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work was supported by the Spanish Ministry of Science and Innovation PGC2018-094765-B-I00 project (MCIN/AEI/10.13039/501100011033/ FEDER “Una manera de hacer Europa”); the project PID2021-122255NB-I00 supported by MCIN/AEI/10.13039/501100011033/ FEDER, UE; the María de Maeztu Center of Excellence CEX2021-001159-M (supported by MCIN/AEI/10.13039/501100011033); the 2021SGR-00356 Consolidated Research Group of the Catalan Government, and the ICREA Acadèmia Distinguished Professorship awarded to CE.

## References

- Abboub, N., Nazzi, T., and Gervain, J. (2016). Prosodic grouping at birth. *Brain Lang.* 162, 46–59. doi: 10.1016/j.bandl.2016.08.002
- Aiken, S. J., and Picton, T. W. (2008). Envelope and spectral frequency-following responses to vowel sounds. *Hear. Res.* 245, 35–47. doi: 10.1016/j.heares.2008.08.004
- Anbuhl, K. L., Uhler, K. M., Werner, L. A., and Tollin, D. J. (2016). “Early development of the human auditory system” in *Fetal and neonatal physiology*. eds. R. A. Polin, S. H. Abman, D. Rowitch and W. E. Benitz. 5th ed (Amsterdam: Elsevier), 1396–1410.
- Andreeva, B., Demenko, G., Möbius, B., Zimmerer, F., Jügler, J., and Jastrzebska, M. (2014). Differences of pitch profiles in Germanic and Slavic languages. In *Proceedings of Interspeech*. ISCA.
- Arenillas-Alcón, S., Costa-Faidella, J., Ribas-Prats, T., Gómez-Roig, M. D., and Escera, C. (2021). Neural encoding of voice pitch and formant structure at birth as revealed by frequency-following responses. *Sci. Rep.* 11:6660. doi: 10.1038/s41598-021-85799-x
- Arenillas-Alcón, S., Ribas-Prats, T., Puertollano, M., Mondéjar-Segovia, A., Gómez-Roig, M. D., Costa-Faidella, J., et al. (2023). Prenatal daily musical exposure is associated with enhanced neural representation of speech fundamental frequency: evidence from neonatal frequency-following responses. *Dev. Sci.* 26:e13362. doi: 10.1111/desc.13362
- Banai, K., Hornickel, J., Skoe, E., Nicol, T., Zecker, S., and Kraus, N. (2009). Reading and subcortical auditory function. *Cereb. Cortex* 19, 2699–2707. doi: 10.1093/cercor/bhp024
- Barac, R., Bialystok, E., Castro, D. C., and Sanchez, M. (2014). The cognitive development of young dual language learners: a critical review. *Early Child. Res. Q.* 29, 699–714. doi: 10.1016/j.ecresq.2014.02.003
- Barkat, T., Polley, D., and Hensch, T. (2011). A critical period for auditory thalamocortical connectivity. *Nat. Neurosci.* 14, 1189–1194. doi: 10.1038/nn.2882
- Basu, M., Krishnan, A., and Weber-Fox, C. (2010). Brainstem correlates of temporal auditory processing in children with specific language impairment. *Dev. Sci.* 13, 77–91. doi: 10.1111/j.1467-7687.2009.00849.x
- Benavides-Varela, S., Hochmann, J. R., Macagno, F., Nespor, M., and Mehler, J. (2012). Newborn's brain activity signals the origin of word memories. *Proc. Natl. Acad. Sci. USA* 109, 17908–17913. doi: 10.1073/pnas.1205413109
- Bialystok, E. (2017). The bilingual adaptation: how minds accommodate experience. *Psychol. Bull.* 143, 233–262. doi: 10.1037/bul0000099
- Boersma, P., and Weenink, D. (2020). *Praat: Doing phonetics by computer* (version 6.1.09). Available at: <http://www.praat.org/>
- Bosch, L., and Sebastián-Gallés, N. (2003). Simultaneous bilingualism and the perception of a language-specific vowel contrast in the first year of life. *Lang. Speech* 46, 217–243. doi: 10.1177/00238309030460020801
- Bosch, L., and Sebastián-Gallés, N. (2010). Evidence of early language discrimination abilities in infants from bilingual environments. *Infancy* 2, 29–49. doi: 10.1207/S15327078IN0201\_3
- Byers-Heinlein, K., Burns, T. C., and Werker, J. F. (2010). The roots of bilingualism in newborns. *Psychol. Sci.* 21, 343–348. doi: 10.1177/0956797609360758
- Byers-Heinlein, K., Espósito, A. G., Winsler, A., Marian, V., Castro, D. C., and Luk, G. (2019). The case for measuring and reporting bilingualism in developmental research. *Collabra Psychol.* 5:37. doi: 10.1525/collabra.233
- Cabrera, L., and Gervain, J. (2020). Speech perception at birth: the brain encodes fast and slow temporal information. *Sci. Adv.* 6:eaba 7830. doi: 10.1126/sciadv.aba7830
- Carcagno, S., and Plack, C. J. (2017). “Short-term learning and memory: training and perceptual learning” in *The frequency-following response: A window into human communication*, vol. 61. eds. N. Kraus, S. Anderson, T. White-Schwoch, R. R. Fay and A. N. Popper (London: Springer Nature), 75–100.

## Acknowledgments

The authors would like to thank to all families who selflessly participated in this study and made it possible.

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The author(s) declared that they were an editorial board member of *Frontiers*, at the time of submission. This had no impact on the peer review process and the final decision.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## Supplementary material

The Supplementary material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fnhum.2024.1379660/full#supplementary-material>

- Chandrasekaran, B., Hornickel, J., Skoe, E., Nicol, T., and Kraus, N. (2009). Context-dependent encoding in the human auditory brainstem relates to hearing speech in noise: implications for developmental dyslexia. *Neuron* 64, 311–319. doi: 10.1016/j.neuron.2009.10.006
- Christophe, A., Mehler, J., and Sebastian-Galles, N. (2001). Perception of prosodic boundary correlates by newborn infants. *Infancy* 2, 385–394. doi: 10.1207/S15327078IN0203\_6
- Conboy, B. T., and Kuhl, P. K. (2011). Impact of second-language experience in infancy: brain measures of first- and second-language speech perception. *Dev. Sci.* 14, 242–248. doi: 10.1111/j.1467-7687.2010.00973.x
- Cooley, J. W., and Tukey, J. W. (1965). An algorithm for the machine calculation of complex Fourier series. *Math. Comput.* 19, 297–301. doi: 10.1090/S0025-5718-1965-0178586-1
- Costa, A., Hernández, M., and Sebastián-Gallés, N. (2008). Bilingualism aids conflict resolution: evidence from the ANT task. *Cognition* 106, 59–86. doi: 10.1016/j.cognition.2006.12.013
- DeCasper, A. J., and Fifer, W. (1980). Of human bonding: newborns prefer their mothers' voice. *Science* 208, 1174–1176. doi: 10.1126/science.7375928
- DeCasper, A. J., and Spence, M. J. (1986). Prenatal maternal speech influences newborns' perception of speech sounds. *Infant Behav. Dev.* 9, 133–150. doi: 10.1016/0163-6383(86)90025-1
- Deutsch, D., Le, J., Shen, J., and Henthorn, T. (2009). The pitch levels of female speech in two Chinese villages. *J. Acoust. Soc. Am.* 125, EL208–EL213. doi: 10.1121/1.3113892
- Eggermont, J. J. (2001). Between sound and perception: reviewing the search for a neural code. *Hear. Res.* 157, 1–42. doi: 10.1016/S0378-5955(01)00259-3
- Font-Alaminos, M., Cornella, M., Costa-Faidella, J., Hervás, A., Leung, S., Rueda, I., et al. (2020). Increased subcortical neural responses to repeating auditory stimulation in children with autism spectrum disorder. *Biol. Psychol.* 149:107807. doi: 10.1016/j.biopsycho.2019.107807
- Gerhardt, K. J., and Abrams, R. M. (2000). Fetal exposures to sound and vibroacoustic stimulation. *J. Perinatol.* 20, S21–S30. doi: 10.1038/sj.jp.7200446
- Gervain, J. (2015). Plasticity in early language acquisition: the effects of prenatal and early childhood experience. *Curr. Opin. Neurobiol.* 35, 13–20. doi: 10.1016/j.conb.2015.05.004
- Gervain, J. (2018). The role of prenatal experience in language development. *Curr. Opin. Behav. Sci.* 21, 62–67. doi: 10.1016/j.cobeha.2018.02.004
- Gervain, J., and Mehler, J. (2010). Speech perception and language acquisition in the first year of life. *Annu. Rev. Psychol.* 61, 191–218. doi: 10.1146/annurev.psych.093008.100408
- Gervain, J., and Werker, J. F. (2008). How infant speech perception contributes to language acquisition. *Lang. Linguist. Compass* 2, 1149–1170. doi: 10.1111/j.1749-818X.2008.00089.x
- Gorina-Careta, N., Ribas-Prats, T., Arenillas-Alcón, S., Puertollano, M., Gómez-Roig, M. D., and Escera, C. (2022). Neonatal frequency-following responses: a methodological framework for clinical applications. *Semin. Hear.* 43, 162–176. doi: 10.1055/s-0042-1756162
- Gorina-Careta, N., Ribas-Prats, T., Costa-Faidella, J., and Escera, C. (2019). "Auditory frequency-following responses" in *Encyclopedia of computational neuroscience*. eds. D. Jaeger and R. Jung (New York, NY: Springer), 1–13.
- Granier-Deferre, C., Ribeiro, A., Jacquet, A. Y., and Bassereau, S. (2011). Near-term fetuses process temporal features of speech. *Dev. Sci.* 14, 336–352. doi: 10.1111/j.1467-7687.2010.00978.x
- Hammer, C. S., Hoff, E., Uchikoshi, Y., Gillanders, C., Castro, D. C., and Sandilos, L. E. (2014). The language and literacy development of young dual language learners: a critical review. *Early Child. Res. Q.* 29, 715–733. doi: 10.1016/j.ecresq.2014.05.008
- Hoff, E. (2003). The specificity of environmental influence: socioeconomic status affects early vocabulary development via maternal speech. *Child Dev.* 74, 1368–1378. doi: 10.1111/1467-8624.00612
- Hornickel, J., Anderson, S., Skoe, E., Yi, H.-G., and Kraus, N. (2012). Subcortical representation of speech fine structure relates to reading ability. *Neuro Report* 23, 6–9. doi: 10.1097/WNR.0b013e32834d2ff
- Järvinen, K., Laukkanen, A.-M., and Aaltonen, O. (2013). Speaking a foreign language and its effect on F0. *Logoped. Phoniatr. Vocol.* 38, 47–51. doi: 10.3109/14015439.2012.687764
- Jeng, F. C., Hu, J., Dickman, B., Montgomery-Reagan, K., Tong, M., Wu, G., et al. (2011). Cross-linguistic comparison of frequency-following responses to voice pitch in American and Chinese neonates and adults. *Ear Hear.* 32, 699–707. doi: 10.1097/AUD.0b013e31821cc0df
- Jeng, F. C., Lin, C.-D., and Wang, T.-C. (2016). Subcortical neural representation to mandarin pitch contours in American and Chinese newborns. *J. Acoust. Soc. Am.* 139, EL190–EL195. doi: 10.1121/1.4953998
- Jeng, F. C., Schnabel, E. A., Dickman, B. M., Hu, J., Li, X., Lin, C.-D., et al. (2010). Early maturation of frequency-following responses to voice pitch in infants with normal hearing. *Percept. Mot. Skills* 111, 765–784. doi: 10.2466/10.22.24.PMS.111.6.765-784
- Joint Committee on Infant Hearing (2019). Year 2019 position statement: principles and guidelines for early hearing detection and intervention programs. *Pediatrics* 106, 798–817. doi: 10.1542/peds.106.4.798
- Keating, P., and Kuo, G. (2012). Comparison of speaking fundamental frequency in English and mandarin. *J. Acoust. Soc. Am.* 132, 1050–1060. doi: 10.1121/1.4730893
- King, C., Warrier, C. M., Hayes, E., and Kraus, N. (2002). Deficits in auditory brainstem pathway encoding of speech sounds in children with learning problems. *Neurosci. Lett.* 319, 111–115. doi: 10.1016/S0304-3940(01)02556-3
- Kovács, Á. M., and Mehler, J. (2009). Flexible learning of multiple speech structures in bilingual infants. *Science* 325, 611–612. doi: 10.1126/science.1173947
- Kraus, N., and Chandrasekaran, B. (2010). Music training for the development of auditory skills. *Nat. Rev. Neurosci.* 11, 599–605. doi: 10.1038/nrn2882
- Krizman, J., and Kraus, N. (2019). Analyzing the FFR: a tutorial for decoding the richness of auditory function. *Hear. Res.* 382, 107779–107174. doi: 10.1016/j.heares.2019.107779
- Krizman, J., Marian, V., Shook, A., Skoe, E., and Kraus, N. (2012). Subcortical encoding of sound is enhanced in bilinguals and relates to executive function advantages. *PNAS* 109, 7877–7881. doi: 10.1073/pnas.1201575109
- Krizman, J., Skoe, E., Marian, V., and Kraus, N. (2014). Bilingualism increases neural response consistency and attentional control: evidence for sensory and cognitive coupling. *Brain Lang.* 128, 34–40. doi: 10.1016/j.bandl.2013.11.006
- Krizman, J., Slater, J., Skoe, E., Marian, V., and Kraus, N. (2015). Neural processing of speech in children is influenced by bilingual experience. *Neurosci. Lett.* 0, 48–53. doi: 10.1016/j.neulet.2014.11.011.Neural
- Kroll, J. F., Dussias, P. E., Bogulski, C. A., and Valdes Kroff, J. R. (2012). Chapter seven – Juggling two languages in one mind: what bilinguals tell us about language processing and its consequences for cognition. *Psychol. Learn. Motiv.* 56, 229–262. doi: 10.1016/B978-0-12-394393-4.00007-8
- Kuhl, P. K. (2010). Brain mechanisms in early language acquisition. *Neuron* 67, 713–727. doi: 10.1016/j.neuron.2010.08.038
- Kuhl, P. K., Stevens, E., Hayashi, A., Deguchi, T., Kiritani, S., and Iverson, P. (2006). Infants show a facilitation effect for native language phonetic perception between 6 and 12 months. *Dev. Sci.* 9, F13–F21. doi: 10.1111/j.1467-7687.2006.00468.x
- Kuhl, P. K., Stevenson, J., Corrigan, N. M., Van den Bosch, J. J. F., Deniz Can, D., and Richards, T. (2016). Neuroimaging of the bilingual brain: structural brain correlates of listening and speaking in a second language. *Brain Lang.* 162, 1–9. doi: 10.1016/j.bandl.2016.07.004
- Lam, S. S.-Y., White-Schwoch, T., Zecker, S. G., Hornickel, J., and Kraus, N. (2017). Neural stability: a reflection of automaticity in reading. *Neuropsychologia* 103, 162–167. doi: 10.1016/j.neuropsychologia.2017.07.023
- Li, P., Legault, J., and Litcofsky, K. A. (2014). Neuroplasticity as a function of second language learning: anatomical changes in the human brain. *Cortex* 58, 301–324. doi: 10.1016/j.cortex.2014.05.001
- Liu, F., Maggu, A. R., Lau, J. C. Y., and Wong, P. C. M. (2015). Brainstem encoding of speech and musical stimuli in congenital amusia: evidence from Cantonese speakers. *Front. Hum. Neurosci.* 8, 1–19. doi: 10.3389/fnhum.2014.01029
- Luk, G. (2017). "Bilingualism" in *The Cambridge encyclopedia of child development*. eds. B. Hopkins, E. Geangu and S. Linkenauer. 2nd ed (Cambridge: Cambridge University Press), 385–391.
- Majewski, W., Hollien, H., and Zalewski, J. (1972). Speaking fundamental frequency of polish adult males. *Phonetica* 25, 119–125. doi: 10.1159/000259375
- 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. doi: 10.1016/j.cub.2009.09.064
- Marian, V., Blumenfeld, H. K., and Kaushanskaya, M. (2007). The language experience and proficiency questionnaire (LEAP-Q): assessing language profiles in bilinguals and multilinguals. *J. Speech Lang. Hear. Res.* 50, 940–967. doi: 10.1044/1092-4388(2007)067
- Mariani, B., Nicoletti, G., Barzon, G., Ortiz-Barajas, M. C., Shukla, M., Guevara, R., et al. (2023). Prenatal experience with language shapes the brain. *Sci. Adv.* 9:eadj3524. doi: 10.1126/sciadv.adj3524
- Marquina Zazura, M. (2011). *Estudio acústico de la variación inter e intralocutor en la frecuencia fundamental de hablantes bilingües de catalán y de castellano* [Universitat Autònoma de Barcelona]. Available at: <https://ddd.uab.cat/record/77033>
- Mary, L., and Yegnanarayana, B. (2008). Extraction and representation of prosodic features for language and speaker recognition. *Speech Comm.* 50, 782–796. doi: 10.1016/j.specom.2008.04.010
- May, L., Byers-Heinlein, K., Gervain, J., and Werker, J. F. (2011). Language and the newborn brain: does prenatal language experience shape the neonate neural response to speech? *Front. Psychol.* 2, 1–9. doi: 10.3389/fpsyg.2011.00222
- Molnar, M., Lallier, M., and Carreiras, M. (2014). The amount of language exposure determines nonlinguistic tone grouping biases in infants from a bilingual environment. *Lang. Learn.* 64, 45–64. doi: 10.1111/lang.12069
- Moon, C., Cooper, R. P., and Fifer, W. P. (1993). Two-day-olds prefer their native language. *Infant Behav. Dev.* 16, 495–500. doi: 10.1016/0163-6383(93)80007-U
- Moon, C., and Fifer, W. (2000). Evidence of transnatal auditory learning. *J. Perinatol.* 20, S37–S44. doi: 10.1038/sj.jp.7200448

- Moon, C., Lagercrantz, H., and Kuhl, P. K. (2012). Language experienced in utero affects vowel perception after birth: a two-country study. *Acta Paediatr.* 102, 156–160. doi: 10.1111/apa.12098
- Moore, J. K., and Linthicum, F. H. (2007). The human auditory system: a timeline of development. *Int. J. Audiol.* 46, 460–478. doi: 10.1080/14992020701383019
- Muscaccia, G., Sams, M., Skoe, E., and Kraus, N. (2007). Musicians have enhanced subcortical auditory and audiovisual processing of speech and music. *Proc. Natl. Acad. Sci. USA* 104, 15894–15898. doi: 10.1073/pnas.0701498104
- 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, 756–766. doi: 10.1037/0096-1523.24.3.756
- Ordin, M., and Mennen, I. (2017). Cross-linguistic differences in bilinguals' fundamental frequency ranges. *J. Speech Lang. Hear. Res.* 60, 1493–1506. doi: 10.1044/2016\_JSLHR-S-16-0315
- Otto-Meyer, S., Krizman, J., White-Schwoch, T., and Kraus, N. (2018). Children with autism spectrum disorder have unstable neural responses to sound. *Exp. Brain Res.* 236, 733–743. doi: 10.1007/s00221-017-5164-4
- Pallier, C., Bosch, L., and Sebastián-Gallés, N. (1997). A limit on behavioral plasticity in vowel acquisition. *Cognition* 64, B9–B17. doi: 10.1016/S0010-0277(97)00030-9
- Pallier, C., Colomé, A., and Sebastián-Gallés, N. (2001). The influence of native-language phonology on lexical access: exemplar-based versus abstract lexical entries. *Psychol. Sci.* 12, 445–449. doi: 10.1111/1467-9280.00383
- Partanen, E., Kujala, T., Näätänen, R., Liitola, A., Sambeth, A., and Huotilainen, M. (2013a). Learning-induced neural plasticity of speech processing before birth. *PNAS* 110, 15145–15150. doi: 10.1073/pnas.1302159110
- Partanen, E., Kujala, T., Tervaniemi, M., and Huotilainen, M. (2013b). Prenatal music exposure induces long-term neural effects. *PLoS One* 8:e78946. doi: 10.1371/journal.pone.0078946
- Partanen, E., Mårtensson, G., Hugoson, P., Huotilainen, M., Fellman, V., and Ådén, U. (2022). Auditory processing of the brain is enhanced by parental singing for preterm infants. *Front. Neurosci.* 16:772008. doi: 10.3389/fnins.2022.772008
- Passoni, E., De Leeuw, E., and Levon, E. (2022). Bilinguals produce pitch range differently in their two languages to convey social meaning. *Lang. Speech* 65, 1071–1095. doi: 10.1177/00238309221105210
- Plack, C. J., Barker, D., and Hall, D. A. (2014). Pitch coding and pitch processing in the human brain. *Hear. Res.* 307, 53–64. doi: 10.1016/j.heares.2013.07.020
- Ramus, F., Hauser, M., Miller, C., Morris, D., and Mehler, J. (2000). Language discrimination by human newborns and by cotton-top tamarin monkeys. *Science* 288, 349–351. doi: 10.1126/science.288.5464.349
- Ressel, V., Pallier, C., Ventura-Campos, N., Díaz, B., Roessler, A., Ávila, C., et al. (2012). An effect of bilingualism on the auditory cortex. *J. Neurosci.* 32, 16597–16601. doi: 10.1523/JNEUROSCI.1996-12.2012
- Ribas-Prats, T., Almeida, L., Costa-Faidella, J., Plana, M., Corral, M. J., Gómez-Roig, M. D., et al. (2019). The frequency-following response (FFR) to speech stimuli: a normative dataset in healthy newborns. *Hear. Res.* 371, 28–39. doi: 10.1016/j.heares.2018.11.001
- Ribas-Prats, T., Arenillas-Alcón, S., Pérez-Cruz, M., Costa-Faidella, J., Gómez-Roig, M. D., and Escera, C. (2023). Speech-encoding deficits in neonates born large-for-gestational age as revealed with the frequency-following response. *Ear Hear.* 44, 829–841. doi: 10.1097/AUD.0000000000001330
- Ribas-Prats, T., Arenillas-Alcón, S., Lip-Sosa, D. L., Costa-Faidella, J., Mazarico, E., Gómez-Roig, M. D., et al. (2021). Deficient neural encoding of speech sounds in term neonates born after fetal growth restriction. *Dev. Sci.* 25:e13189. doi: 10.1111/desc.13189
- Rosenthal, M. A. (2020). A systematic review of the voice-tagging hypothesis of speech-in-noise perception. *Neuropsychologia* 136:107256. doi: 10.1016/j.neuropsychologia.2019.107256
- Rowe, M. (2008). Child-directed speech: relation to socioeconomic status, knowledge of child development and child vocabulary skill. *J. Child Lang.* 35, 185–205. doi: 10.1017/S0305000907008343
- Ruben, R. J. (1995). The ontogeny of human hearing. *Int. J. Pediatr. Otorhinolaryngol.* 32, S199–S204. doi: 10.1016/0165-5876(94)01159-U
- Russo, N. M., Nicol, T. G., Zecker, S. G., Hayes, E. A., and Kraus, N. (2005). Auditory training improves neural timing in the human brainstem. *Behav. Brain Res.* 156, 95–103. doi: 10.1016/j.bbr.2004.05.012
- Saffran, J. R., Werker, J. F., and Werner, L. A. (2006). “The Infant’s auditory world: hearing, speech, and the beginnings of language” in *Handbook of child psychology: cognition, perception, and language*. eds. D. Kuhn, R. S. Siegler, W. Damon and R. M. Lerner (Hoboken, NJ: John Wiley & Sons, Inc), 58–108.
- Sansavini, A., Bertoncini, J., and Giovanelli, G. (1997). Newborns discriminate the rhythm of multisyllabic stressed words. *Dev. Psychol.* 33, 3–11. doi: 10.1037/0012-1649.33.1.3
- Schochat, E., Rocha-Muniz, C. N., and Filippini, R. (2017). “Understanding auditory processing disorder through the FFR” in *The frequency-following response: A window into human communication*. eds. N. Kraus, S. Anderson, T. White-Schwoch, R. Fay and A. Popper (London: Springer International Publishing), 225–250.
- Sebastián-Gallés, N., Rodríguez-Fornells, A., de Diego-Balaguer, R., and Díaz, B. (2006). First- and second-language phonological representations in the mental lexicon. *J. Cogn. Neurosci.* 18, 1277–1291. doi: 10.1162/jocn.2006.18.8.1277
- Skoe, E., Burakiewicz, E., Figueiredo, M., and Hardin, M. (2017). Basic neural processing of sound in adults is influenced by bilingual experience. *Neuroscience* 349, 278–290. doi: 10.1016/j.neuroscience.2017.02.049
- Skoe, E., and Kraus, N. (2010). Auditory brain stem response to complex sounds: a tutorial. *Ear Hear.* 31, 302–324. doi: 10.1097/AUD.0b013e3181c8db272
- Song, J. H., Skoe, E., Wong, P. C. M., and Kraus, N. (2008). Plasticity in the adult human auditory brainstem following short-term linguistic training. *J. Cogn. Neurosci.* 20, 1892–1902. doi: 10.1162/jocn.2008.20.131
- Sundara, M., Polka, L., and Molnar, M. (2008). Development of coronal stop perception: bilingual infants keep pace with their monolingual peers. *Cognition* 108, 232–242. doi: 10.1016/j.cognition.2007.12.013
- The Jamovi Project. (2023). *Jamovi 2.3*. Available at: <https://www.jamovi.org>.
- The Mathworks Inc. (2019). *MATLAB R2019b*, Natick, Massachusetts.
- Van Bezooijen, R. (1995). Sociocultural aspects of pitch differences between Japanese and Dutch women. *Lang. Speech* 38, 253–265. doi: 10.1177/002383099503800303
- Vouloumanos, A., and Werker, J. F. (2007). Listening to language at birth: evidence for a bias for speech in neonates. *Dev. Sci.* 10, 159–164. doi: 10.1111/j.1467-7687.2007.00549.x
- Weaver, I., Cervoni, N., Champagne, F., D’Alessio, A., Sharma, S., Seckl, J., et al. (2004). Epigenetic programming by maternal behavior. *Nat. Neurosci.* 7, 847–854. doi: 10.1038/nn1276
- Werker, J. F., and Curtin, S. (2005). PRIMIR: a developmental framework of infant speech processing. *Lang. Learn. Dev.* 1, 197–234. doi: 10.1080/15475441.2005.9684216
- Werker, J. F., and Hensch, T. (2015). Critical periods in speech perception: new directions. *Annu. Rev. Psychol.* 66, 173–196. doi: 10.1146/annurev-psych-010814-015104
- Werker, J. F., and Tees, R. C. (2005). Speech perception as a window for understanding plasticity and commitment in language systems of the brain. *Dev. Psychobiol.* 46, 233–251. doi: 10.1002/dev.20060
- White-Schwoch, T., Davies, E. C., Thompson, E. C., Woodruff Carr, K., Nicol, T., Bradlow, A. R., et al. (2015). Auditory-neurophysiological responses to speech during early childhood: effects of background noise. *Hear. Res.* 328, 34–47. doi: 10.1016/j.heares.2015.06.009