

1 **Longitudinal trajectories of the neural encoding mechanisms of**
2 **speech-sound features during the first year of life**

3 Marta Puertollano^{a,b,c}, Teresa Ribas-Prats^{a,b,c}, Natàlia Gorina-Careta^{a,b,c}, Siham Ijjou-
4 Kadiri^{a,b,c}, Sonia Arenillas-Alcón^{a,b,c}, Alejandro Mondéjar-Segovia^{a,b,c}, María Dolores
5 Gómez-Roig^{c,d}, Carles Escera^{a,b,c}

6
7 **Affiliations:**

8 ^aBrainlab – Cognitive Neuroscience Research Group. Department of Clinical Psychology and
9 Psychobiology, University of Barcelona (Catalonia, Spain)

10 ^bInstitute of Neurosciences, University of Barcelona (Catalonia, Spain)

11 ^cInstitut de Recerca Sant Joan de Déu, Esplugues de Llobregat (Barcelona, Spain)

12 ^dBCNatal – Barcelona Center for Maternal Fetal and Neonatal Medicine (Hospital Sant Joan de
13 Déu and Hospital Clínic), University of Barcelona (Catalonia, Spain)

14
15 **Corresponding authors:**

16 Carles Escera, PhD, Professor. Department of Clinical Psychology and Psychobiology,
17 University of Barcelona. Passeig de la Vall d'Hebron 171, 08035 Barcelona (Catalonia-
18 Spain). Tel: +34 933 125 048, Fax: +34 934 021 584. E-mail: cescera@ub.edu

19
20
21
22 **Conflict of Interest Disclosure:** The authors have no conflict of interest to declare.

23 **Data Availability Statement:** The data that support the findings of this study are
24 available from the corresponding author Carles Escera (cescera@ub.edu) upon
25 reasonable request.

26 **Funding:** This work was supported by the Spanish Ministry of Science and Innovation
27 projects PGC2018-094765-B-I00 [MCIN/AEI/10.13039/501100011033/ FEDER “Una
28 manera de hacer Europa”] and PID2021-122255NB-I00 [MCIN/ AEI /
29 10.13039/501100011033 / FEDER, UE], the María de Maeztu Center of Excellence
30 [CEX2021-001159-M by
31 MCIN/AEI/10.13039/501100011033MCIN/AEI/10.13039/501100011033], and the
32 ICREA Acadèmia Distinguished Professorship awarded to Carles Escera.

33 **Longitudinal trajectories of the neural encoding mechanisms of speech-sound**
34 **features during the first year of life**

35 **Abstract**

36 Infants learn to recognize the sounds of their mother language very early in
37 development, exhibiting a remarkable ability to process the fine-grained
38 spectrotemporal characteristics of speech. However, the neural machinery underlying
39 this ability is not yet understood. Here, we used an auditory evoked potential termed
40 frequency-following response (FFR) to unravel the maturational pattern of the neural
41 encoding of two speech sound characteristics: voice pitch and temporal fine structure,
42 during the first year of life. The FFR was elicited to a two-vowel stimulus (/oa/) in a
43 sample of 41 healthy-term neonates that were tested at birth and retested at the ages of
44 six and twelve months. Results revealed a shortening of neural phase-locking time to
45 stimulus envelope from birth to six months, with no further difference between six and
46 twelve months. While neural encoding of voice pitch was similar across age, encoding
47 of the stimulus temporal fine structure exhibited a rapid maturation from birth to six
48 months, without further improvement from this age onwards. Results highlight the first
49 six months of age as a crucial period in the maturation of the neural encoding
50 mechanisms of the temporal fine structure of speech sounds, essential for phoneme
51 discrimination during early language acquisition.

52

53 **Keywords:** Early language acquisition, speech encoding, infant, newborn, Frequency-
54 Following Response, auditory evoked potential

55 **1. Introduction**

56 Infants show a native talent for language acquisition even since the very early stages of
57 development. Behavioral evidence has shown that infants follow similar developmental
58 trajectories in the acquisition of their mother tongue regardless of their culture and
59 language, evolving from babbling (5 to 10 months; Kuhl, 2004) to their first words
60 utterances by the age of 10 to 18 months (Feldman, 2019). Yet, there is a lack of
61 consensus regarding the neural mechanisms supporting this talent. Starting from the
62 tuning to the phonetic repertoire of the mother tongue, language acquisition entails a
63 sophisticated fine-grained neural machinery across the entire auditory system to capture
64 the complex spectro-temporal acoustic features that characterize the speech sounds.

65 Despite postnatal hearing experience is essential for an adequate auditory and language
66 development, neonates are born with a wide range of universal speech perception
67 abilities that allow them to acquire any language. For instance, neonates can
68 discriminate different languages they have not been exposed to if these are rhythmically
69 different (Byers-Heinlein et al., 2010; Mehler et al., 1988; Nazzi et al., 1998). They can
70 also encode the pitch of a speech sound in an adult-like manner (Arenillas-Alcón et al.,
71 2021; Jeng et al., 2011, 2016), as well as recognize their mothers' voice (Decasper &
72 Fifer, 1980; Hepper et al., 1993) and melodies they have been exposed to during
73 pregnancy (Granier-Deferre et al., 2011).

74 Yet to imitate speech and acquire their mother tongue, babies still need to learn to
75 classify non-identical sounds into either separate or equivalent phonetic categories in a

76 language dependent manner. This competence emerges a few months later after birth, in
77 parallel with the myelination trajectory of the auditory pathway (Moore et al., 1995) and
78 the exposure to their specific native language (Kuhl et al., 1992, 2003; Rivera-Gaxiola
79 et al., 2005), a process mediated by statistical learning (Jusczy et al., 1994). Previous
80 studies using behavioral paradigms demonstrated that by the age of six months, babies
81 are able to perceive the variability inherent in each phonetic unit and thus to identify
82 and discern vowels that are typical of their mother tongue, which alters phonetic
83 perception towards a native-like model (Kuhl et al., 1992; Maye et al., 2002). Speech
84 perception at this stage, as measured following the conditioned head turn paradigm,
85 may serve as a predictor of future infant language abilities at the age of 13, 16 and 24
86 months (Tsao et al., 2004). A decline in the discrimination ability for nonnative
87 contrasts in language is evident by the end of the first year of life (Rivera-Gaxiola et al.,
88 2005; Tsao et al., 2006; Werker and Tees, 1984), when infants start to develop an adult-
89 like attunement to their native language phoneme repertoire (Cheour et al., 1998; Kuhl
90 et al., 2006; Werker et al., 1981).

91 Research on neural mechanisms underlying the acquisition of speech sounds has
92 benefited from advances in the use of neonatal and infant brain potentials. One such
93 evoked potential is the mismatch negativity (MMN), elicited to acoustic and linguistic
94 changes (Kujala et al., 2023). Another recently growing body of research is exploring
95 the so-called frequency-following response (FFR). The FFR is a non-invasive auditory
96 evoked potential that is elicited to periodic complex stimuli such as speech or music,

197 and reflects neural activity from the auditory nerve to the cortex. It mimics the eliciting
198 stimulus, thus providing a unique snapshot into the neural encoding of the two
199 distinctive features that characterize the speech sounds: voice pitch, associated to its
200 fundamental frequency, and the temporal fine structure, associated to its formants
201 (Aiken & Picton, 2008; Coffey et al., 2019; Gorina-Careta et al., 2021; Krizman &
202 Kraus, 2019). The FFR has been studied in infancy to characterize normal and abnormal
203 developmental trajectories of neural speech encoding (Banai et al., 2005, 2009;
204 Cunningham et al., 2001), as these appear disrupted in children with literacy
205 impairments, including dyslexia (Hornickel & Kraus, 2013), specific language
206 impairment (Basu et al., 2010), and autism (Font-Alaminos et al., 2020; Otto-Meyer et
207 al., 2018; Russo et al., 2008).

208 The FFR has also been explored during the first months of life in several cross-sectional
209 studies with different age periods of interest, as an attempt to describe its typical
210 trajectory along early development. A decrease in neural conduction times and neural
211 phase-locking onset has been already observed at the early age of 45 days (Ferreira et
212 al., 2021), with further shortenings until the age of ten months (Anderson et al., 2015).
213 An adult-like voice pitch encoding at birth has also been reported (Arenillas-Alcón et
214 al., 2021; Jeng et al., 2016), with a more robust neural representation with age across
215 the first year of life (Jeng et al., 2010; Ribas-Prats et al., in press; Van Dyke et al.,
216 2017). The maturation of neural encoding of temporal fine structure components, as
217 assessed through neural responses to the high-frequency formants and harmonics,

118 begins as early as the first month of life (Ribas-Prats et al., in press) and continuous to
119 develop until the age of ten months (Anderson et al., 2015).

120 However, the studies reviewed above provide an incomplete view of the developmental
121 trajectory of speech-sounds neural encoding mechanisms during the first year of age.
122 Behavioral paradigms impose constraints on disentangling the neural underpinnings of
123 speech perception. Furthermore, cross-sectional designs adopted by previous
124 electrophysiological studies offer a limited approach to characterize the neural
125 correlates of speech development. The present longitudinal study was set to provide a
126 pioneering and comprehensive picture of the maturational pattern of the neural
127 mechanisms involved in encoding two distinct speech-sound features during the first
128 postnatal year, as reflected in the FFR: voice pitch, as represented by its fundamental
129 frequency, and speech temporal fine structure, corresponding to its formants. We
130 hypothesized an enhancement in the neural encoding of these two speech-sound features
131 as a function of age, starting from birth to six months and further continuing from six
132 months to the age of one year. Neural phase-locking onset was also expected to
133 decrease due to the well-known myelination process of the auditory pathway during the
134 first year of life.

135 **2. Methods**

136 **2.1 Participants**

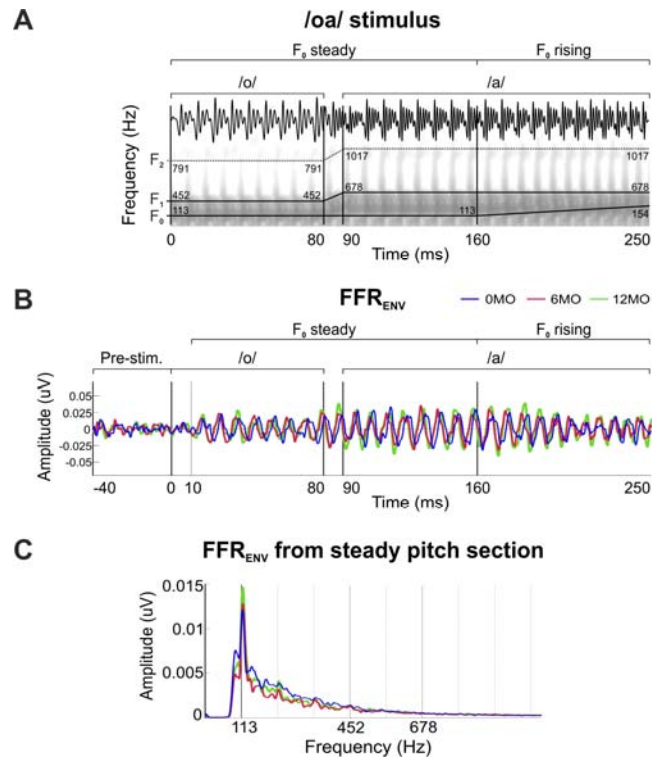
137 Sixty-six healthy-term neonates were recruited at the SJD Barcelona Children's
138 Hospital (Catalonia, Spain). Forty-one of them (21 females; mean age = 1.75 ± 1.09
139 days after birth; mean gestational age = 39.71 ± 0.91 weeks) were followed-up in all
140 stages of the study: at the age of six months (aged 5.53 to 7.5 months after birth; mean =
141 6.41 ± 0.37 SD) and at twelve months of age (aged 11.97 to 13.7 months after birth;
142 mean = 12.60 ± 0.40 SD). All neonates were born after low-risk gestations, without
143 either pathologies or risk factors for hearing impairment (Joint Committee on Infant
144 Hearing, 2019). Apgar scores were higher than 7 at 1 and 5 minutes after birth and, in
145 all cases, birth weight was adequate for their gestational age (Figueras & Gratacós,
146 2014). Furthermore, all infants had passed the universal hearing screening test as part of
147 the standard medical routine, based on an automated auditory brainstem response
148 system to ensure auditory pathway health (ALGO 3i, Natus Medical Incorporated, San
149 Carlos, CA). To confirm the integrity of the auditory pathway, an auditory brainstem
150 response (ABR) to a click stimulus (10 μ s; delivered monaurally to the right ear at 60
151 dB SPL at a rate of 19.30 Hz, for a total of 4000 averaged sweeps) was also obtained
152 from every neonate.

153 The study was approved by the Bioethics Committee of SJD Barcelona Children's
154 Hospital (Internal review board ID: PIC-185-19). A written informed consent was
155 obtained from all parents or legal guardians prior to the data collection in accordance
156 with the Code of Ethics of the World Medical Association (Declaration of Helsinki).

157 The data that supports the findings of this study and the code used for data analysis are
158 available upon reasonable request to the corresponding author.

159 **2.2 Stimulus**

160 To obtain the FFR, a two-vowel /oa/ stimulus was used (see Fig. 1A), as previously
161 designed in our laboratory by Arenillas-Alcón et al. (2021). The stimulus had a total
162 duration of 250 ms, with a steady F_0 at 113 Hz for its first 160 ms and a linearly rising
163 F_0 from 113 to 154 Hz during its last 90 ms (from 160 to 250 ms). The stimulus was
164 presented monaurally to the right ear at a rate of 3.39 Hz and an intensity of 60 dB SPL
165 in alternating polarities through Etymotic shielded earphones of 300 Ω (ER, Elk Grove
166 Village, IL, EEUU) connected to a Flexicoupler® disposable adaptor (Natus Medical
167 Incorporated, San Carlos, CA).



168

169 **Fig. 1.** (A) Temporal and spectral plots of the /oa/ stimulus with a schematic representation of
170 its formant structure. F₀, F₁ and F₂ are defined for the different stimulus sections. (B) Grand-
171 averaged FFR_{ENV} waveforms in the time domain from the 41 infants recorded at birth (0MO,
172 blue), six months (6MO, red) and twelve months of age (12MO, green), obtained after averaged
173 neural response polarities. (C) Amplitude FFR_{ENV} spectra extracted from the analyzed steady
174 part of the stimulus (10-160 ms).

175 2.3 Procedure

176 After a successful universal hearing screening test, the ABR and subsequently the FFR
177 were recorded at the hospital room while the newborns were sleeping at their crib,
178 following the same protocol used in previous studies (see Arenillas-Alcón et al., 2021,
179 2023; Ribas-Prats et al., 2019, 2022, 2023a, in press; for a review see Gorina-Careta et
180 al., 2022). Recording was interrupted to any sign of discomfort or sleep disruption and it

181 was resumed when the newborn was asleep again. Infants that successfully completed
182 the neonatal FFR session were invited to two successive follow-up sessions: at six and
183 twelve months of age. Out of the 68 recruited neonates only 41 completed the two
184 follow-up sessions, yielding a failure to complete the longitudinal study at one year in
185 39.7% of the participants. The infants that returned were retested at six and at twelve
186 months of age at a hospital dispensary, keeping the baby either asleep or as calm as
187 possible during the recording in order to ensure the highest quality of the EEG data. The
188 total mean duration of the sessions was around 30 minutes, including a preparation time
189 of around 5 minutes, 20 minutes of recording (four /oa/ stimulus blocks \times 1000 sweeps
190 \times 295 ms SOA), and the additional time for the rejected sweeps.

191 **2.4 Data acquisition**

192 ABR and FFR recordings were carried out using a SmartEP platform connected to a
193 Duet amplifier, including the cABR and Advanced Hearing Research modules
194 (Intelligent Hearing Systems, Miami, FL, USA). Three disposable Ag/AgCl electrodes
195 located in a vertical montage were employed for the recording, with the active electrode
196 placed at Fpz, ground at the forehead and reference at the right mastoid. Impedances
197 were kept below 10 k Ω for all electrodes. The continuous FFR signal was acquired at a
198 sampling rate of 13333 Hz with an online bandpass filter between 30 and 1500 Hz.
199 Online data was epoched from -40.95 (for the baseline period) to 249.975 ms (relative
200 to the stimulus onset). An automated online rejection of artifacts was applied, excluding
201 any sweep with voltage values exceeding $\pm 30 \mu\text{V}$.

202 **2.5 Data processing**

203 An offline bandpass filter from 80 to 1500 Hz was applied for FFR analysis. Neural
204 responses to alternating polarities were averaged $[(\text{Condensation} + \text{Rarefaction})/2]$ to
205 emphasize the FFR components related to the stimulus envelope (FFR_{ENV}) and to
206 minimize the involvement of putative cochlear microphonics. In addition, to assess the
207 neural encoding of the vowels' first formant (F_1) and minimizing the contribution of
208 envelope related activity (Aiken & Picton, 2008; Krizman & Kraus, 2019), the FFR
209 temporal fine structure (FFR_{TFS}) was analyzed by subtracting the neural response to the
210 alternating polarities $[(\text{Rarefaction} - \text{Condensation})/2]$. Only the steady part of the
211 stimulus was considered for the analysis. Thus, the FFR_{TFS} (spectral peaks
212 corresponding to F_1) was analyzed separately for the /o/ section (10 to 80 ms, $F_0 = 113$
213 Hz, $F_1 = 452$ Hz) and for the /a/ steady section (90 to 160 ms, $F_0 = 113$ Hz, $F_1 = 678$
214 Hz).

215 FFR parameters were evaluated using custom scripts from Matlab R2019b (The
216 Mathworks Inc., 2019) used in previous studies performed in our laboratory (Arenillas-
217 Alcón et al., 2021; Ribas-Prats et al., 2019). A comprehensive description is given
218 below for the three parameters extracted and tested separately for the different
219 frequencies of interest (for a detailed description, see Arenillas-Alcon et al., 2021;
220 Ribas-Prats et al., 2019).

221 **Neural lag.** In order to assess the neural transmission delay occurring along the auditory
222 pathway, the neural lag value was obtained. This parameter accounts for the time lag

223 between the stimulus presentation and the neural phase-locking onset (Arenillas-Alcón
224 et al., 2021; Jeng et al., 2010; Ribas-Prats et al., 2019). It was computed from the cross-
225 correlation between the stimulus and the neural response as the time shift that
226 corresponded to the maximum cross-correlation magnitude.

227 ***Spectral amplitude.*** In order to analyze the neural-phase locking magnitude at the
228 frequency of interest (F_0 , 113 Hz; /o/ F_1 , 452 Hz; /a/ F_1 , 678 Hz), spectral amplitude was
229 obtained as an indicator of the response strength at that given frequency (Arenillas-
230 Alcón et al., 2021; Ribas-Prats et al., 2019; White-Schwoch et al., 2015b). To obtain the
231 FFR frequency decomposition, spectral amplitude was calculated after applying the Fast
232 Fourier Transform (FFT; Cooley & Tukey, 1965), by computing the mean amplitude
233 within a ± 5 Hz frequency window centered at the frequency peak of interest. Spectral
234 amplitude at F_0 was retrieved from the FFR_{ENV} corresponding to the /oa/ steady section
235 (10 to 160 ms) to assess voice pitch encoding of the speech-sound stimulus. Spectral
236 amplitudes at the stimulus F_1 frequencies were extracted separately from the FFR_{TFS}
237 corresponding to the /o/ section (10 to 80 ms) and the /a/ steady section (90 to 160 ms).

238 ***Signal-to-noise ratio.*** Signal-to-noise ratio (SNR) at the frequency peak of interest was
239 calculated in order to estimate the FFR relative spectral magnitude. It was computed by
240 dividing the spectral amplitude value obtained for the given frequency of interest (± 5
241 Hz window centered at the peak of interest) by the mean amplitude of its two flanks (28
242 Hz windows centered at ± 19 Hz from the frequency of interest). SNR at F_0 was
243 extracted from the FFR_{ENV} to evaluate voice pitch encoding. SNRs at vowels F_1 were

244 retrieved from the FFR_{TFS} to assess the formant structure encoding of the auditory
245 stimulus and analyzed following the same procedure as for the spectral amplitude
246 parameter (i.e., the values were extracted separately from the neural responses to the
247 vowel sections).

248 **2.6 Statistical analysis**

249 Statistical analyses were performed using SPSS 25.0 (IBM Corp, Armonk, NY).
250 Descriptive statistics are presented for each parameter as median and interquartile range
251 for each time of measurement (see Table 1). Results were considered statistically
252 significant when $p < .05$. Normality was assessed with Shapiro-Wilk's test and, as all
253 parameters followed a non-normal distribution, Friedman's test was applied. After a
254 given significant result, Wilcoxon signed-rank test was employed to explore each time
255 point measurements pair.

256 In addition, to ensure that the spectral amplitude and SNR measurements obtained for
257 the stimulus F_1 were specific to the corresponding stimulus vowel section (i.e., 452 Hz
258 for the /o/ vowel, and 678 Hz for the /a/ vowel), as well as its possible interaction with
259 age, a repeated measures Analysis of Variance (rmANOVA) test was performed. For
260 that, the variables Age (0, 6 and 12 months) and Stimulus Section (/o/ and /a/) were
261 chosen as within-subject factors. Bonferroni correction was applied to adjust p -values
262 for multiple pairwise comparisons. Greenhouse-Geisser correction was used when the
263 assumption of sphericity was violated. Partial eta squared (η^2) was reported as a
264 measure of effect size.

265 **Table 1**

266 Descriptive statistics expressed as median (IQR, interquartile range), and Friedman test
 267 comparison between the 41 neonates recorded at birth (0-MO) and their retest at the age of six
 268 (6-MO) and twelve months (12-MO) for each FFR parameter assessed. Wave V amplitude and
 269 latency values at birth are also depicted for the extended 66 neonatal sample as mean (SD).

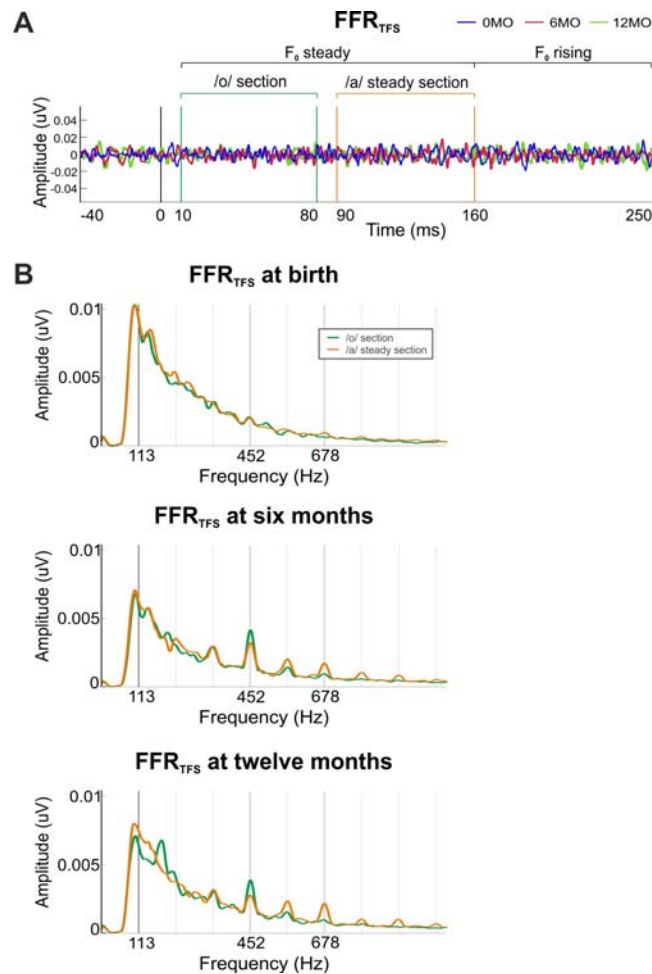
Measures N = 41	0-MO	6-MO	12-MO	Friedman test	df	p value
Wave V						
Amplitude (μ V)	0.08 (0.14)	-	-	-	-	-
Latency (ms)	8.50 (0.38)	-	-	-	-	-
FFR						
Neural lag (ms)	8.03 (2.03)	6.68 (1.95)	6.53 (1.76)	22.84	40	<.001
FFR_{ENV}						
Spectral amplitude F ₀ (μ V)	.008 (.008)	.009 (.012)	.013 (.014)	2.10	-	.35
SNR F ₀ (dB)	3.99 (5.69)	4.75 (7.90)	6.41 (5.41)	1.81	-	.41
FFR_{TFS}						
Spectral amplitude at /o/ F ₁ (10-80 ms; μ V)	.0018 (.002)	.0031 (.005)	.0027 (.005)	4.44	-	.10
SNR at /o/ F ₁ (10-90ms; dB)	1.38 (4.88)	4.83 (6.98)	3.52 (4.96)	15.85	-	<.001
Spectral amplitude at /a/ F ₁ (90-160 ms; μ V)	.0008 (.001)	.0012 (.002)	.0017 (.002)	6.78	-	.034
SNR at /a/ F ₁ (90-160 ms; dB)	1.73 (4.08)	2.61 (5.99)	3.81 (3.79)	7.37	-	.025

270 **3. Results**

271 An ABR to a click stimulus was obtained from every neonate before the FFR recording
 272 to confirm the integrity of the auditory pathway. All recruited infants (N = 66) had an
 273 identifiable wave V peak at birth, with a mean latency of 8.50 (\pm 0.38 SD) ms and a
 274 mean amplitude of 0.08 (\pm 0.14 SD) μ V (Table 1). Values were similar to those
 275 previously reported at the literature (Arenillas et al., 2021; Ribas-Prats et al., 2019;
 276 Stuart et al., 1994).

277 In order to unravel the maturational pattern of the neural encoding of speech sounds
 278 features during the first year of life, FFRs elicited to the /oa/ stimulus were collected
 279 from the sample of forty-one neonates that completed the follow-up at the ages of six
 280 and twelve months. The corresponding grand-average FFR_{ENV} and FFR_{TFS} waveforms

281 are shown in Fig. 1B and Fig. 2A respectively. Table 1 depicts the descriptive statistics
282 and results from the Friedman test comparison for all FFR parameters evaluated at the
283 three developmental stages (i.e., 0, 6 and 12 months).

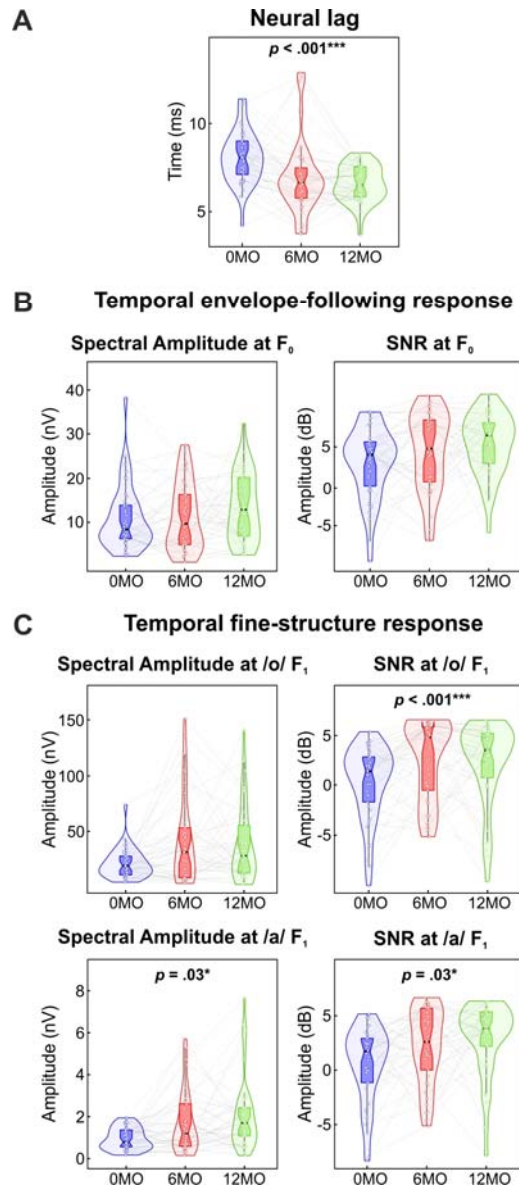


284

285 **Fig. 2.** (A) Time-domain grand-averaged FFR_{TFS} waveforms extracted after subtracting neural
286 responses to alternating stimulus polarities from the 41 infants recorded at birth (0MO, blue)
287 and retested at six months (6MO, red) and twelve months of age (12MO, green). (B) Amplitude
288 FFR_{TFS} spectra obtained for the two vowel sections: /o/ (green), /a/ (orange).

289 **3.1 Neural lag**

290 Results revealed a consistently shortened neural phase-locking onset as a function of
291 age ($X^2(2) = 27.84, p < .001$; see Fig. 3A). Post-hoc analyses revealed shorter neural lag
292 at both six ($Mdn = 6.68; z = -3.51, p < .001$) and twelve months of age ($Mdn = 6.53; z =$
293 $-4.39, p < .001$) in comparison to that at birth ($Mdn = 8.03$). Neural transmission delay
294 at the age of six and twelve months were similar ($z = -.97, p = .33$). To further support
295 the results obtained, statistical analyses were repeated for the entire sample that
296 completed the follow-up session at six months of age ($N = 54$; 27 females; mean age at
297 birth = 1.81 ± 1.28 days; mean gestational age = 39.73 ± 0.97 weeks; age at six months
298 = 5.53 to 7.77 months after birth; mean age at six months = 6.42 ± 0.43 months).
299 Similar results were obtained for neural transmission delay (i.e., shortened neural lag at
300 six months; $z = -3.79, p < .001$; see Table 2).



301

302 **Fig. 3.** Data distribution from the 41 infants recorded at birth (0MO, blue) and retested at six
 303 (6MO, red) and twelve months of age (12MO, green). Violin plots are depicted for (A) the
 304 neural lag, (B) the FFR_{ENV} obtained to the steady part of the stimulus (10-160 ms), and (C) the
 305 FFR_{TFS} corresponding to /o/ (upper panel) and /a/ (lower panel) stimulus sections.

306 **Table 2**

307 Descriptive statistics and comparison between the sample of 54 neonates recorded at birth (0-
308 MO) and their retest at the age of six (6-MO) for each FFR parameter assessed.

FFR measures N = 54	0-MO	6-MO	<i>t</i> test	<i>df</i>	<i>p</i> value
Neural lag (ms)	8.03 (2.08) ^a	6.64 (1.82) ^a	-3.79 ^b	53	<.001
FFR _{ENV}					
Spectral amplitude F ₀ (μV)	.008 (.008) ^a	.009 (.010) ^a	-0.28 ^b	-	.78
SNR F ₀ (dB)	3.12 (3.98)	4.27 (4.99)	-1.55	-	.12
FFR _{TFS}					
Spectral amplitude at /o/ F ₁ (10-80 ms; μV)	.0017 (.002) ^a	.0032 (.005) ^a	-3.72 ^b	-	<.001
SNR at /o/ F ₁ (10-90ms; dB)	1.03 (5.12) ^a	4.38 (6.67) ^a	-3.24 ^b	-	.001
Spectral amplitude at /a/ F ₁ (90-160 ms; μV)	.0008 (.001) ^a	.0012 (.002) ^a	-2.97 ^b	-	.003
SNR at /a/ F ₁ (90-160 ms; dB)	1.63 (5.15) ^a	2.65 (5.70) ^a	-2.31 ^b	-	.021

309 Results are expressed as mean (SD). ^a Median (IQR, interquartile range). ^b Wilcoxon Signed-Rank Test.

310 3.2 Temporal envelope-following response

311 Grand-average FFR_{ENV} waveforms were obtained at each developmental stage (see Fig.
312 1B). The corresponding frequency spectrum for averaged polarities is shown in Fig. 1C.
313 The strength of the stimulus F₀ neural representation along the first year of life was
314 assessed by means of the spectral amplitude and SNR parameters computed on the
315 FFR_{ENV}. Statistical analyses revealed no significant differences across age (at birth, at
316 six and at twelve months) in neither spectral amplitude parameter ($X^2(2) = 2.10$, $p =$
317 $.35$) or in SNR ($X^2(2) = 1.81$, $p = .41$; see Fig. 3B). Results remained statistically
318 similar for both parameters in the analysis with the extended fifty-four infants sample
319 that could complete the recording at the age of six months (i.e., spectral amplitude, $z = -$
320 $.28$, $p = .78$; SNR, $t_{(53)} = -1.55$, $p = .13$).

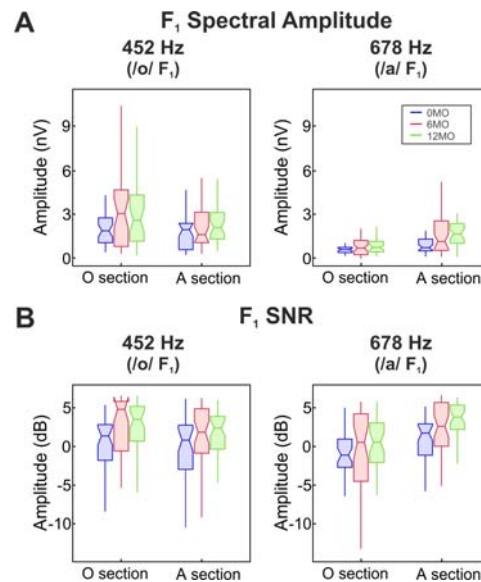
321 3.3 Temporal fine-structure response

322 The maturation of neural mechanisms for the encoding of the speech-sound formant
323 structure along the first year of life was analyzed from the FFR_{TFS}. Grand-average
324 FFR_{TFS} waveforms are illustrated in Fig. 2A for each developmental stage. In order to
325 evaluate phase-locking at the stimulus F₁, neural responses to each stimulus vowel
326 section were assessed separately. Fig. 2B illustrates the frequency spectrum
327 corresponding to both vowel sections at birth, six and twelve months of age. Spectral
328 amplitudes and SNRs were retrieved selecting the spectral peaks corresponding to the
329 frequency of interest according to each vowel (452 Hz for the /o/; 678 Hz for the /a/).
330 Fig. 3C depicts spectral amplitude and SNR values for the FFR_{TFS} along the three
331 developmental stages.

332 ***Neural encoding of the /o/ vowel F₁.*** Spectral amplitude and SNR at /o/ vowel F₁ (452
333 Hz) were analyzed at the corresponding stimulus /o/ vowel section. No differences were
334 observed for spectral amplitude between the three stages of development ($X^2(2) = 4.44$,
335 $p = .1$; see Table 1). For the SNR, significant differences were obtained ($X^2(2) = 15.85$,
336 $p < .001$), with larger values at six months ($Mdn = 4.83$; $z = -2.73$, $p = .006$) and twelve
337 months of age ($Mdn = 3.52$; $z = -2.60$, $p = .009$) in comparison to birth ($Mdn = 1.38$);
338 SNR values at six and twelve months of age were similar ($z = -.72$, $p = .47$). Wilcoxon
339 signed-rank test assessed for the extended six-months old sample (i.e., fifty-four infants)
340 revealed larger values at six months of age in the two parameters assessed (spectral
341 amplitude, $z = -3.72$, $p < .001$; SNR, $z = -3.24$, $p = .001$; see Table 2).

342 To investigate the specificity of the neural encoding of the formant structure
343 corresponding to each of the two vowels of the /oa/ stimulus, and its possible interaction
344 with age, a two-way rmANOVA test was conducted with the factors Age (0, 6 and 12
345 months) and Stimulus Section (/o/ and /a/) on the spectral amplitude and its SNR at 452
346 Hz, corresponding to the /o/ F_1 (see Fig. 4). Spectral amplitude results revealed a main
347 effect of stimulus section ($F_{(1,40)} = 7.96, p = .007, \eta^2 = .17$), with higher spectral
348 amplitudes for the /o/ section ($M = .0033 \pm < .001$) compared to the /a/ section ($M =$
349 $.0026 \pm < .001$). A main effect of age was also observed ($F_{(2,80)} = 4.50, p = .014, \eta^2 =$
350 $.10$), with larger amplitudes at the age of six months ($M = .004 \pm .001; p = .012$) and the
351 age of twelve months ($M = .003 \pm < .001; p = .048$) in comparison to birth ($M = .002 \pm$
352 $< .001$). No significant developmental changes between six and twelve months-old
353 stages were found ($p = 1$). Interaction between age and stimulus section was not
354 significant ($F_{(2,80)} = 2.95, p = .06, \eta^2 = .07$).

355 Significant differences in SNR were not observed for the stimulus section ($F_{(1,40)} =$
356 $2.70, p = .108, \eta^2 = .06$). A main effect of age ($F_{(2,80)} = 8.10, p = .001, \eta^2 = .17$) was
357 obtained with the same pattern as that observed for spectral amplitude. Larger SNRs
358 were found at both six months ($M = 1.99 \pm .57; p = .009$) and twelve months of age (M
359 $= 2.10 \pm .41; p = .001$) in comparison to birth ($M = -.13 \pm .53$); with no significant
360 variation from six to twelve months of age ($p = 1$). Significant differences were not
361 found for the age per stimulus section interaction ($F_{(2,80)} = .19, p = .825, \eta^2 = .005$).



362

363 **Fig. 4.** Neural encoding of the first formant corresponding to the two vowel steady sections
364 measured in (A) spectral amplitude and (B) SNR. Data are plotted from infants at birth (blue),
365 six months (red) and twelve months of age (green), and for /o/ F₁ (452 Hz, left) and /a/ F₁ (678
366 Hz, right) encoding at the different stages of development.

367 **Neural encoding of the /a/ vowel F₁.** In order to estimate the encoding of the /a/ vowel

368 F₁ (678 Hz), spectral amplitudes and SNRs were evaluated for the /a/ vowel section.

369 Significant differences were found as a function of age for spectral amplitude ($X^2(2) =$

370 6.78, $p = .034$), with larger values at six months ($Mdn = .0012$; $z = -2.74$, $p = .006$) and

371 twelve months of age ($Mdn = .0017$; $z = -4.88$, $p < .001$) in comparison to values

372 obtained at birth ($Mdn = .0008$). Remarkably, spectral amplitudes obtained at twelve

373 months were also statistically larger than the those at six months ($z = -2.15$, $p = .032$).

374 Similar results were observed for SNR ($X^2(2) = 7.37$, $p = .025$), where infants presented

375 larger SNRs at the age of six months ($Mdn = 2.61$; $z = -2.12$, $p = .034$) and at twelve

376 months of age ($Mdn = 3.81$; $z = -2.84$, $p = .004$) compared to birth ($Mdn = 1.73$). No

377 differences in SNR values were observed between six and twelve months of age ($z = -$

378 1.07, $p = .29$). Wilcoxon signed-rank test assessed for the extended six-months old
379 sample (i.e., fifty-four infants) also indicated larger values at six months of age than at
380 birth in the two parameters assessed (spectral amplitude, $z = -2.67$, $p = .003$; SNR, $z = -$
381 2.31 , $p = .021$).

382 Two-way rmANOVA tests were hence conducted to examine neural response
383 specificity to Stimulus Section (/o/ and /a/) as a function of Age (0, 6 and 12 months) on
384 the spectral amplitude and its SNR at 678 Hz, corresponding to the /a/ F_1 (see Fig. 4).
385 Spectral amplitude yielded a main effect of stimulus section ($F_{(1,40)} = 36.36$, $p < .001$,
386 $\eta^2 = .48$), with higher values for the /a/ section ($M = .002 \pm <.001$) in comparison to
387 the /o/ section ($M = .001 \pm <.001$). A main effect of age was also revealed ($F_{(2,80)} = 7.92$,
388 $p = .001$, $\eta^2 = .17$), with significantly larger values at both six months ($M = .0013 \pm$
389 $<.001$; $p = .010$) and twelve months of age ($M = .0015 \pm <.001$; $p = .001$) in comparison
390 to the ones obtained at the moment of birth ($M = .0008 \pm <.001$); but no significant
391 changes appeared from six to twelve months age ($p = .89$). The age per stimulus section
392 interaction was significant ($F_{(1,74,69,50)} = 5.99$, $p = .006$, $\epsilon = .83$). *Post-hoc* analysis
393 revealed higher spectral amplitudes at 678 Hz during the /a/ vs. the /o/ vowel sections at
394 the three stages of development (birth, $z = -3.40$, $p = .001$; six-months, $z = -3.34$, $p =$
395 $.001$; twelve-months, $z = -4.32$, $p < .001$).

396 SNR results yielded a main effect of stimulus section ($F_{(1,40)} = 38.90$, $p < .001$, $\eta^2 =$
397 $.49$), indicating higher values at the /a/ section ($M = 1.93 \pm .31$) in comparison to the /o/
398 section ($M = -.39 \pm .30$). A main effect of age ($F_{(2,80)} = 3.83$, $p = .026$, $\eta^2 = .09$), with

399 significantly larger values at twelve months of age ($M = 1.68 \pm .42$) than at birth ($M = -$
400 $.18 \pm .43$; $p = .019$). No significant age per stimulus section interaction was found
401 ($F_{(1.72,69.06)} = .80$, $p = .44$, $\epsilon = .86$).

402 **4. Discussion**

403 This study describes the longitudinal trajectory of speech-sound neural encoding
404 mechanisms required for language acquisition across the first year of life, as examined
405 through the electrophysiological recordings of the frequency-following response (FFR)
406 elicited to the /oa/ syllable at birth, six and twelve months of age. While no age-related
407 changes were observed in the encoding of voice pitch, a significant enhancement was
408 depicted across the first six-month postnatal period in neural transmission times and
409 neural encoding of the stimulus temporal fine structure. Results contribute to knowledge
410 from previous studies on the developmental trajectory of speech-sound neural encoding
411 mechanisms (Anderson et al., 2015; Arenillas-Alcón et al., 2021, Ribas-Prats et al.
412 2019, in press), by specifically filling the gap with a longitudinal sample spanning the
413 first year of life, and by unveiling an early neural maturation in the neural encoding of
414 the speech temporal fine structure. Results point to a sensitive developmental window in
415 the emergence of core neural mechanisms required for speech acquisition that occurs
416 within the first six months after birth, which may underlie the co-occurring critical
417 behavioral language milestones (i.e., acquisition of phonetic categories).

418 Language acquisition relies on an accurate development of the auditory brain, which is
419 already functional to process sounds at the beginning of the third trimester of pregnancy

420 (Hepper & Shahidullah, 1994; Moore & Linthicum, 2007; Querleu et al., 1988; Ruben,
421 1995). Around the 27th gestational week, the first traces of myelin can be observed in
422 both the cochlear nerve (Moore & Lithicum, 2001) and the brainstem auditory pathway
423 (Moore, et al., 1995), paralleling the first fetuses' behavioral and electrophysiological
424 outcomes to auditory stimulation (Draganova et al., 2018; Hepper & Shahidullah, 1994;
425 Schneider et al., 2001). At birth, the cochlea has reached its adult size and is fully
426 functional (Lavigne-Rebillard & Dan Bagger-Sjöbäck, 1992; Moore & Linthicum,
427 2007), but the auditory brain is not completely mature yet. The shortened neural
428 transmission delay observed in our results at the age of six months aligns with prior
429 literature showing acceleration of auditory neural responses as a function of
430 development (Amorim et al., 2009; Anderson et al., 2015; Madrid et al., 2021; Ribas-
431 Prats et al., in press; Sharma et al., 2016). This decrease in neural lag can be attributed
432 to age-related increasing myelination in the auditory white matter tracts along the
433 brainstem, midbrain (Moore et al., 1995), and primary auditory cortex (Su et al., 2008)
434 that occurs during this early period of development.

435 The perception of speech sound F_0 and F_1 plays a crucial role in language acquisition, as
436 these acoustic features constitute key cues that facilitate the learning of a native
437 language (Moon & Hong, 2014). Indeed, voice pitch perception is defined by sound F_0
438 (Oxenham, 2012) and provides phonological, syntactic and semantic cues needed for
439 distinguishing words units from a continuous speech (Nakatani & Schaffer, 1978) and
440 for speaker identification (Mary & Yegnanarayana, 2008). Early sensory experience in

441 utero has been demonstrated to be a prerequisite for auditory learning and neural
442 plasticity during the perinatal period (Draganova et al., 2018; Webb et al., 2015). Once
443 they are born, neonates can track the sound envelope, irrespective of the language they
444 have been exposed to during pregnancy (Ortiz Barajas et al., 2021). Our results support
445 previous findings describing an intrauterine perception of low-frequency speech cues
446 (Hepper & Shahidullah, 1994; Voegtline et al., 2013) and an adult-like voice pitch
447 encoding at birth (Anderson et al., 2015; Arenillas-Alcón et al., 2021; Jeng et al., 2011),
448 as no age-related changes were observed here in neither spectral amplitude and SNR
449 values at the stimulus F_0 peak.

450 Some discrepancies emerge in literature in relation to the developmental pattern of
451 speech F_0 neural encoding during the first year of life. While most of the studies report a
452 more robust neural encoding of pitch as a function of age, this pattern does not
453 constantly reach statistical significance across the literature. For instance, Jeng and
454 colleagues (2010) found pitch encoding improvement on a single infant tested at
455 different time-points through the first ten months of age (i.e. 1, 3, 5, 7 and 10 months).
456 Similarly, Van Dyke et al. (2017) described stronger F_0 neural encoding when
457 comparing a group of older infants (7-12 months old) with a group of younger infants
458 (2-7 months old). Ribas-Prats and colleagues (in press) also observed age-related
459 improvements from the first postnatal month to six months of age in a longitudinal
460 sample of healthy-term neonates. Yet, in a cross-sectional study performed by Anderson
461 et al. (2015), a similar but not significant trend was found in F_0 neural encoding for a

462 sample of infants aged from 3 to 10 months. Similarly, our results show a linear pattern
463 of stronger pitch encoding through age during the first postnatal year, although this
464 increase did not reach statistical significance. These discrepancies may arise from the
465 inconsistency on the stimuli frequency components employed across various studies or
466 the individual linguistic environment to which individuals from different study samples
467 are exposed, as suggested by the linguistic experience model (see Jeng et al., 2011;
468 Kuhl et al., 1992). Notably, Jeng and colleagues (2011) compared two samples of
469 neonates exposed to different languages during pregnancy (i.e., English and Chinese)
470 with two matched samples of adults of the same native languages. While Chinese adults
471 showed larger pitch strength values compared to their matched neonates, pitch strength
472 values were comparable between the American neonatal and adult samples.

473 Stimulus F_1 is closely linked to the discrimination of vowel sounds (Kieft et al., 2010,
474 2013; Nenadić et al., 2020) and phoneme recognition (Diehl and Lindblom, 2004).
475 Auditory postnatal experience is essential for infants to encode high-frequency
476 components of speech sounds. During pregnancy, the maternal womb acts as a low-pass
477 filter and limits auditory stimulation as it attenuates frequencies above 500 Hz
478 (Gerhardt & Abrahms, 1996, 2000; Hepper & Shahidullah, 1994; Parga et al., 2018),
479 which impedes neonates from hearing high-frequency components before birth. The
480 higher spectral amplitude and SNR values observed at the /a/ vowel F_1 peak (i.e. 678
481 Hz) as a function of age support a non-mature neural encoding at birth of these
482 frequency components above circa 500 Hz. Similarly, the ability to track the formant

483 structure of speech seems to be not fully developed at birth, but postnatally experience-
484 dependent, as supported by the higher SNR at both vowels' F_1 by the age of six months.
485 These results align with the spectrally ascendant developmental pattern of the auditory
486 system described by Graven and Browne (2008), stating that low-frequency sounds are
487 tuned first in the cochlea, and highlighting the period from 25 gestational weeks to six
488 months of age as the most critical in the neurosensory development of the auditory
489 system. Moreover, the absence of further differences in SNR values at both vowels F_1
490 between six and twelve months of age highlights a special relevance of the first six-
491 months postnatal period on the maturation of the temporal fine structure encoding of
492 speech. These results support previous cross-sectional FFR studies reporting an
493 enhancement of F_1 neural encoding as a function of age (Anderson et al., 2015; Van
494 Dyke et al., 2017), as well as the longitudinal findings by Ribas-Prats et al. (in press),
495 but further extend these latter findings on the longitudinal trajectory on neural F_1
496 encoding through early development to the age of twelve months.

497 This language-specific attunement to frequencies across the auditory pathway is
498 essential for early language acquisition, as it relies on infants' ability to apprehend the
499 phonological structure corresponding to a given language (Best et al., 2016; Cutler,
500 2008). This experience will indeed contribute and facilitate the appropriate
501 identification of native language phonemes by the age of six months (Best et al., 2016;
502 Cheour et al., 1998; Kuhl et al., 1992). Moreover, a perceptual re-organization by the
503 age of six months has been previously proposed, suggesting an attentional shift from

504 syllabic units at birth to phonemic units at six months, cues that are more relevant for
505 word and grammar learning (Nallet & Gervain, 2021; Ortiz Barajas et al., 2021). A rich
506 extrauterine auditory environment is key in the improvement of phonemic categories
507 perception in early language acquisition. During this postnatal period, there is a notable
508 increase in social interactions that play a vital role on infant development, such as in the
509 initial coordination of gestures, vocalizations and facial expressions in interaction with
510 others (Kuhl, 2004, 2010). The co-occurring temporal fine structure encoding bootstrap
511 by the early age of six months revealed in our results suggest an outstanding relevance
512 of this novel and rich extrauterine environment on acoustic and language acquisition.

513 Future studies are needed to replicate this pattern of development during the first year of
514 life in healthy-term infants. Our results uncover the first six-months of life as a key
515 period in neural speech encoding development. Thus, it is essential to include earlier
516 developmental stages in future longitudinal studies to fully examine this early period of
517 development. Several language-related disabilities such as dyslexia (Banai et al., 2009),
518 learning related disorders (King et al., 2002) or autism (Russo et al., 2008) have been
519 associated with alterations in the spectro-temporal encoding of complex sounds. Also,
520 clinical conditions that occur during the gestational period such as fetal alcohol
521 syndrome (Wyper & Rasmussen, 2011) or fetal growth restriction (Partanen et al., 2018;
522 Ribas-Prats et al., 2022) have an impact on cognitive outcomes, being language one of
523 the major areas affected. Therefore, research on early maturation of neural speech
524 encoding related to language abnormalities in infancy is the next crucial step to

525 comprehend key differences that underlie an inadequate or delayed language
526 acquisition.

527 Early interventions aimed at improving speech encoding in language affected conditions
528 during the first months of life have not been explored yet. However, the positive effect
529 of musical experience and training in speech encoding has been previously documented
530 for both prenatal (Arenillas-Alcón et al., 2023) and postnatal periods (Wong et al.,
531 2007), suggesting its potential as a promising intervention tool worth exploring. The
532 clinical potential of the electrophysiological tool used in this study has been previously
533 discussed along the literature (Gorina-Careta et al., 2022; Kraus & White-Schwoch,
534 2015a; Ribas-Prats et al., 2019), suggesting the FFR as a potential biomarker of early
535 language acquisition. Using the FFR as a screening test to early detect speech encoding
536 abnormalities could open the possibility to further design and implement preventive
537 protocols for language-related impairments. The present study provides normative FFR
538 values for the first year of life (i.e., at birth and at six and twelve months of age) and it
539 may thus serve as a reference for future studies on speech-sound neural encoding
540 development.

541 **5. Conclusion**

542 The present longitudinal study describes the outstanding maturation of the temporal fine
543 structure neural encoding mechanisms during the very early stages of development. The
544 findings highlight the crucial role of the first six postnatal months in shaping the neural
545 mechanisms that support the encoding of speech sounds, and hence are of major

546 relevance for speech perception and language acquisition. Specifically, our findings
547 unveiled an enhancement in the neural encoding of the formant structure throughout the
548 first six postnatal months, without further maturation up to the first year of life. This
549 reveals a critical maturational period for the neural machinery underlying the ability to
550 discriminate the subtle variations that define phonemes, promoting the formation of
551 phonetic categories. Notably, no significant changes in the neural encoding of voice
552 pitch were observed across this developmental period, which supports a mature voice
553 pitch encoding already at birth. These findings contribute to our understanding of early
554 neural speech encoding and underscore the significance of investigating neural
555 correlates of early speech processing disabilities. Further research in this field can
556 provide valuable guidance for addressing language-related abnormalities and promoting
557 healthy language development in infants.

558 **Acknowledgements**

559 The authors want to express their gratitude to all the families who generously embraced
560 this project for their selfless participation and commitment to the study.

561

562

563 **References**

- 564 Aiken, S. J., & Picton, T. W. (2008). Envelope and spectral frequency-following
565 responses to vowel sounds. *Hearing Research*, 245(1–2), 35–47.
566 <https://doi.org/10.1016/j.heares.2008.08.004>
- 567 American Academy of Pediatrics, Joint Committee on Infant Hearing (2019). Year 2019
568 Position Statement: Principles and Guidelines for Early Hearing Detection and
569 Intervention Programs. *The Journal of Early Hearing Detection and*
570 *Intervention*, 4(2), 1-44. <https://digitalcommons.usu.edu/jehdi/vol4/iss2/1>
- 571 Amorim, R. B., Agostinho-Pesse, R. S., & Alvarenga, K. D. F. (2009). The maturational
572 process of the auditory system in the first year of life characterized by brainstem
573 auditory evoked potentials. *Journal of Applied Oral Science*, 17, 57-62.
574 <https://doi.org/10.1590/S1678-77572009000700010>
- 575 Anderson, S., Parbery-Clark, A., White-Schwoch, T., & Kraus, N. (2015). Development
576 of subcortical speech representation in human infants. *The Journal of the*
577 *Acoustical Society of America*, 137(6), 3346-3355.
578 <https://doi.org/10.1121/1.4921032>
- 579 Arenillas-Alcón, S., Costa-Faidella, J., Ribas-Prats, T., Gómez-Roig, M. D., & Escera,
580 C. (2021). Neural encoding of voice pitch and formant structure at birth as
581 revealed by frequency-following responses. *Scientific reports*, 11(1), 1-16.
582 <https://doi.org/10.1038/s41598-021-85799-x>

- 583 Arenillas-Alcón, S., Ribas-Prats, T., Puertollano, M., Mondéjar-Segovia, A.,
584 Gómez-Roig, M. D., Costa-Faidella, J., & Escera, C. (2023). Prenatal daily
585 musical exposure is associated with enhanced neural representation of speech
586 fundamental frequency: Evidence from neonatal frequency-following
587 responses. *Developmental Science*, 00, e13362.
588 <https://doi.org/10.1111/desc.13362>
- 589 Banai, K., Hornickel, J., Skoe, E., Nicol, T., Zecker, S., & Kraus, N. (2009). Reading
590 and subcortical auditory function. *Cerebral cortex*, 19(11), 2699-2707.
591 <https://doi.org/10.1093/cercor/bhp024>
- 592 Banai, K., Nicol, T., Zecker, S. G., & Kraus, N. (2005). Brainstem timing: implications
593 for cortical processing and literacy. *Journal of Neuroscience*, 25(43), 9850-9857.
594 <https://doi.org/10.1523/JNEUROSCI.2373-05.2005>
- 595 Basu, M., Krishnan, A., & Weber-Fox, C. (2010). Brainstem correlates of temporal
596 auditory processing in children with specific language impairment.
597 *Developmental science*, 13(1), 77-91. [https://doi.org/10.1111/j.1467-](https://doi.org/10.1111/j.1467-7687.2009.00849.x)
598 [7687.2009.00849.x](https://doi.org/10.1111/j.1467-7687.2009.00849.x)
- 599 Best, C. T., Goldstein, L. M., Nam, H., & Tyler, M. D. (2016). Articulating what infants
600 attune to in native speech. *Ecological Psychology*, 28(4), 216-261.
601 <https://doi.org/10.1080/10407413.2016.1230372>

- 602 Byers-Heinlein, K., Burns, T. C., & Werker, J. F. (2010). The roots of bilingualism in
603 newborns. *Psychological science*, *21*(3), 343-348.
604 <https://doi.org/10.1177/0956797609360758>
- 605 Cheour, M., Ceponiene, R., Lehtokoski, A., Luuk, A., Allik, J., Alho, K., & Näätänen,
606 R. (1998). Development of language-specific phoneme representations in the
607 infant brain. *Nature neuroscience*, *1*(5), 351-353. <https://doi.org/10.1038/1561>
- 608 Coffey, E. B., Nicol, T., White-Schwoch, T., Chandrasekaran, B., Krizman, J., Skoe, E.,
609 Zatorre, R. J., & Kraus, N. (2019). Evolving perspectives on the sources of the
610 frequency-following response. *Nature communications*, *10*(1), 5036.
611 <https://doi.org/10.1038/s41467-019-13003-w>
- 612 Cooley, J. W., & Tukey, J. W. (1965). An algorithm for the machine calculation of
613 complex Fourier series. *Mathematics of computation*, *19*(90), 297-301.
614 <https://doi.org/10.2307/2003354>
- 615 Corp., I. SPSS 25.0. (Corp., I, Chicago). <https://www.ibm.com/>
- 616 Cunningham, J., Nicol, T., Zecker, S. G., Bradlow, A., & Kraus, N. (2001).
617 Neurobiologic responses to speech in noise in children with learning problems:
618 deficits and strategies for improvement. *Clinical Neurophysiology*, *112*(5), 758-
619 767. [https://doi.org/10.1016/s1388-2457\(01\)00465-5](https://doi.org/10.1016/s1388-2457(01)00465-5)

- 620 Cutler, A. (2008). The 34th Sir Frederick Bartlett Lecture: The abstract representations
621 in speech processing. *Quarterly journal of experimental psychology*,
622 *61*(11), 1601-1619. <https://doi.org/10.1080/13803390802218542>
- 623 DeCasper, A. J., & Fifer, W. P. (1980). Of human bonding: Newborns prefer their
624 mothers' voices. *Science*, *208*(4448), 1174-1176.
625 <https://doi.org/10.1126/science.7375928>
- 626 Diehl, R.L., Lindblom, B. (2004). Explaining the Structure of Feature and Phoneme
627 Inventories: The Role of Auditory Distinctiveness. In: *Speech Processing in the*
628 *Auditory System*. Springer Handbook of Auditory Research, vol 18. Springer,
629 New York, NY. https://doi.org/10.1007/0-387-21575-1_3
- 630 Draganova, R., Schollbach, A., Schleger, F., Braendle, J., Brucker, S., Abele, H.,
631 Kagan, K. O., Wallwiener, D., Fritsche, A., Eswaran, H., & Preissl, H. (2018).
632 Fetal auditory evoked responses to onset of amplitude modulated sounds. A fetal
633 magnetoencephalography (fMEG) study. *Hearing Research*, *363*, 70-77.
634 <https://doi.org/10.1016/j.heares.2018.03.005>
- 635 Feldman, H. M. (2019). How young children learn language and speech. *Pediatrics in*
636 *review*, *40*(8), 398-411. <https://doi.org/10.1542/pir.2017-0325>
- 637 Ferreira, L., Skarzynski, P. H., Skarzynska, M. B., Sanfins, M. D., & Biaggio, E. P. V.
638 (2021). Effect of Auditory Maturation on the Encoding of a Speech Syllable in

- 639 the First Days of Life. *Brain Sciences*, 11(7), 844.
640 <https://doi.org/10.3390/brainsci11070844>
- 641 Figueras, F., & Gratacós, E. (2014). Update on the diagnosis and classification of fetal
642 growth restriction and proposal of a stage-based management protocol. *Fetal*
643 *diagnosis and therapy*, 36(2), 86-98. <https://doi.org/10.1159/000357592>
- 644 Font-Alaminos, M., Cornella, M., Costa-Faidella, J., Hervás, A., Leung, S., Rueda, I., &
645 Escera, C. (2020). Increased subcortical neural responses to repeating auditory
646 stimulation in children with autism spectrum disorder. *Biological psychology*,
647 149, 107807. <https://doi.org/10.1016/j.biopsycho.2019.107807>
- 648 Gerhardt, K. J., & Abrams, R. M. (1996). Fetal hearing: characterization of the stimulus
649 and response. *Seminars in perinatology*, 20(1), 11–20.
650 [https://doi.org/10.1016/s0146-0005\(96\)80053-x](https://doi.org/10.1016/s0146-0005(96)80053-x)
- 651 Gerhardt, K. J., & Abrams, R. M. (2000). Fetal exposures to sound and vibroacoustic
652 stimulation. *Journal of Perinatology*, 20(1), S21-S30.
653 <https://doi.org/10.1038/sj.jp.7200446>
- 654 Gorina-Careta, N., Kurkela, J. L., Hämäläinen, J., Astikainen, P., & Escera, C. (2021).
655 Neural generators of the frequency-following response elicited to stimuli of low
656 and high frequency: A magnetoencephalographic (MEG) study. *Neuroimage*,
657 231, 117866. <https://doi.org/10.1016/j.neuroimage.2021.117866>

- 658 Gorina-Careta, N., Ribas-Prats, T., Arenillas-Alcón, S., Puertollano, M., Gómez-Roig,
659 M. D., & Escera, C. (2022). Neonatal frequency-following responses: A
660 methodological framework for clinical applications. *Seminars in Hearing*,
661 43(03), 162-176. <https://doi.org/10.1055/s-0042-1756162>
- 662 Granier-Deferre, C., Bassereau, S., Ribeiro, A., Jacquet, A. Y., & DeCasper, A. J.
663 (2011). A melodic contour repeatedly experienced by human near-term fetuses
664 elicits a profound cardiac reaction one month after birth. *PLoS One*, 6(2),
665 e17304. <https://doi.org/10.1371/journal.pone.0017304>
- 666 Graven, S. N., & Browne, J. V. (2008). Auditory development in the fetus and infant.
667 *Newborn and infant nursing reviews*, 8(4), 187-193.
668 <https://doi.org/10.1053/j.nainr.2008.10.010>
- 669 Hepper, P. G., Scott, D., & Shahidullah, S. (1993). Newborn and fetal response to
670 maternal voice. *Journal of Reproductive and Infant Psychology*, 11(3), 147-153.
671 <https://doi.org/10.1080/02646839308403210>
- 672 Hepper, P. G., & Shahidullah, B. S. (1994). The development of fetal hearing. *Fetal and*
673 *Maternal Medicine Review*, 6(3), 167-179.
674 <https://doi.org/10.1017/S0965539500001108>
- 675 Hornickel, J., & Kraus, N. (2013). Unstable representation of sound: a biological marker
676 of dyslexia. *Journal of Neuroscience*, 33(8), 3500-3504.
677 <https://doi.org/10.1523/JNEUROSCI.4205-12.2013>

- 678 Jeng, F. C., Hu, J., Dickman, B., Montgomery-Reagan, K., Tong, M., Wu, G., & Lin, C.
679 D. (2011). Cross-linguistic comparison of frequency-following responses to
680 voice pitch in American and Chinese neonates and adults. *Ear and hearing*,
681 32(6), 699-707. <https://doi.org/10.1097/AUD.0b013e31821cc0df>
- 682 Jeng, F. C., Lin, C. D., & Wang, T. C. (2016). Subcortical neural representation to
683 Mandarin pitch contours in American and Chinese newborns. *The Journal of the*
684 *Acoustical Society of America*, 139(6), EL190.
685 <https://doi.org/10.1121/1.4953998>
- 686 Jeng, F. C., Schnabel, E. A., Dickman, B. M., Hu, J., Li, X., Lin, C. D., & Chung, H. K.
687 (2010). Early maturation of frequency-following responses to voice pitch in
688 infants with normal hearing. *Perceptual and motor skills*, 111(3), 765-784.
689 <https://doi.org/10.2466/10.22.24.PMS.111.6.765-784>
- 690 Jusczyk, P.W., Luce, P.A., & Charles-Luce, J. (1994). Infants' sensitivity to phonotactic
691 patterns in the native language. *Journal of Memory and Language*, 33, 630-645.
692 <https://doi.org/10.1006/jmla.1994.1030>
- 693 Kiefte, M., Enright, T., & Marshall, L. (2010). The role of formant amplitude in the
694 perception of /i/and/u. *The Journal of the Acoustical Society of America*, 127(4),
695 2611-2621. <https://doi.org/10.1121/1.3353124>
- 696 Kiefte, M., Nearey, T. M., Assmann, P. F., Ball, M. J., & Gibbon, F. E. (2013). Vowel
697 perception in normal speakers. *Handbook of vowels and vowel disorders*, 2, 160.

- 698 King, C., Warrier, C. M., Hayes, E., & Kraus, N. (2002). Deficits in auditory brainstem
699 pathway encoding of speech sounds in children with learning problems.
700 *Neuroscience letters*, 319(2), 111-115. [https://doi.org/10.1016/S0304-](https://doi.org/10.1016/S0304-3940(01)02556-3)
701 [3940\(01\)02556-3](https://doi.org/10.1016/S0304-3940(01)02556-3)
- 702 Krizman, J., & Kraus, N. (2019). Analyzing the FFR: A tutorial for decoding the
703 richness of auditory function. *Hearing research*, 382, 107779.
704 <https://doi.org/10.1016/j.heares.2019.107779>
- 705 Kuhl, P. K. (2004). Early language acquisition: cracking the speech code. *Nature*
706 *reviews neuroscience*, 5(11), 831-843. <https://doi.org/10.1038/nrn1533>
- 707 Kuhl, P. K. (2010). Brain mechanisms in early language acquisition. *Neuron*, 67(5),
708 713-727. <https://doi.org/10.1016/j.neuron.2010.08.038>
- 709 Kuhl, P. K., Stevens, E., Hayashi, A., Deguchi, T., Kiritani, S., & Iverson, P. (2006).
710 Infants show a facilitation effect for native language phonetic perception
711 between 6 and 12 months. *Developmental science*, 9(2), F13-F21.
712 <https://doi.org/10.1111/j.1467-7687.2006.00468.x>
- 713 Kuhl, P. K., Tsao, F. M., & Liu, H. M. (2003). Foreign-language experience in infancy:
714 Effects of short-term exposure and social interaction on phonetic learning.
715 *Proceedings of the National Academy of Sciences*, 100(15), 9096-9101.
716 <https://doi.org/10.1073/pnas.1532872100>

- 717 Kuhl, P. K., Williams, K. A., Lacerda, F., Stevens, K. N., & Lindblom, B. (1992).
718 Linguistic experience alters phonetic perception in infants by 6 months of age.
719 *Science*, 255(5044), 606-608. <https://doi.org/10.1126/science.1736364>
- 720 Kujala, T., Partanen, E., Virtala, P., & Winkler, I. (2023). Prerequisites of language
721 acquisition in the newborn brain. *Trends in Neurosciences*.
722 <https://doi.org/10.1016/j.tins.2023.05.011>
- 723 Lavigne-Rebillard, M., & Bagger-Sjöbäck, D. (1992). Development of the human stria
724 vascularis. *Hearing research*, 64(1), 39-51. [https://doi.org/10.1016/0378-
725 5955\(92\)90166-k](https://doi.org/10.1016/0378-5955(92)90166-k)
- 726 Madrid, A. M., Walker, K. A., Smith, S. B., Hood, L. J., & Prieve, B. A. (2021).
727 Relationships between click auditory brainstem response and speech frequency
728 following response with development in infants born preterm. *Hearing
729 Research*, 407, 108277. <https://doi.org/10.1016/j.heares.2021.108277>
- 730 Mary, L., & Yegnanarayana, B. (2008). Extraction and representation of prosodic
731 features for language and speaker recognition. *Speech communication*, 50(10),
732 782-796. <https://doi.org/10.1016/j.specom.2008.04.010>
- 733 Maye, J., Werker, J. F., & Gerken, L. (2002). Infant sensitivity to distributional
734 information can affect phonetic discrimination. *Cognition*, 82(3), B101-B111.
735 [https://doi.org/10.1016/s0010-0277\(01\)00157-3](https://doi.org/10.1016/s0010-0277(01)00157-3)

- 736 Mehler, J., Jusczyk, P., Lambertz, G., Halsted, N., Bertoncini, J., & Amiel-Tison, C.
737 (1988). A precursor of language acquisition in young infants. *Cognition*, 29(2),
738 143-178. [https://doi.org/10.1016/0010-0277\(88\)90035-2](https://doi.org/10.1016/0010-0277(88)90035-2)
- 739 Moon, I. J., & Hong, S. H. (2014). What is temporal fine structure and why is it
740 important?. *Korean journal of audiology*, 18(1), 1.
741 <https://doi.org/10.7874/kja.2014.18.1.1>
- 742 Moore, J. K., & Linthicum Jr, F. H. (2001). Myelination of the human auditory nerve:
743 different time courses for Schwann cell and glial myelin. *Annals of Otology,*
744 *Rhinology & Laryngology*, 110(7), 655-661.
745 <https://doi.org/10.1177/000348940111000711>
- 746 Moore, J. K., & Linthicum Jr, F. H. (2007). The human auditory system: a timeline of
747 development. *International journal of audiology*, 46(9), 460-478.
748 <https://doi.org/10.1080/14992020701383019>
- 749 Moore, J. K., Perazzo, L. M., & Braun, A. (1995). Time course of axonal myelination in
750 the human brainstem auditory pathway. *Hearing Research*, 91(1–2), 208–209.
751 [https://doi.org/10.1016/0378-5955\(95\)00218-9](https://doi.org/10.1016/0378-5955(95)00218-9)
- 752 Nakatani, L. H., & Schaffer, J. A. (1978). Hearing “words” without words: Prosodic
753 cues for word perception. *The Journal of the Acoustical Society of America*,
754 63(1), 234-245. <https://doi.org/10.1121/1.381719>

- 755 Nallet, C., & Gervain, J. (2021). Neurodevelopmental preparedness for language in the
756 neonatal brain. *Annual Review of Developmental Psychology*, 3, 41-58.
757 <https://doi.org/10.1146/annurev-devpsych-050620-025732>
- 758 Nazzi, T., Bertocini, J., & Mehler, J. (1998). Language discrimination by newborns:
759 toward an understanding of the role of rhythm. *Journal of Experimental*
760 *Psychology: Human perception and performance*, 24(3), 756.
761 <https://doi.org/10.1037//0096-1523.24.3.756>
- 762 Nenadić, F., Coulter, P., Nearey, T. M., & Kieft, M. (2020). Perception of vowels with
763 missing formant peaks. *The Journal of the Acoustical Society of*
764 *America*, 148(4), 1911-1921. <https://doi.org/10.1121/10.0002110>
- 765 Ortiz Barajas, M. C., Guevara, R., & Gervain, J. (2021). The origins and development
766 of speech envelope tracking during the first months of life. *Developmental*
767 *cognitive neuroscience*, 48, 100915. <https://doi.org/10.1016/j.dcn.2021.100915>
- 768 Otto-Meyer, S., Krizman, J., White-Schwoch, T., & Kraus, N. (2018). Children with
769 autism spectrum disorder have unstable neural responses to sound. *Experimental*
770 *Brain Research*, 236(3), 733-743. <https://doi.org/10.1007/s00221-017-5164-4>
- 771 Oxenham, A. J. (2012). *Pitch perception*. *Journal of Neuroscience*, 32(39), 13335-
772 13338. <https://doi.org/10.1523/JNEUROSCI.3815-12.2012>
- 773 Parga, J. J., Daland, R., Kesavan, K., Macey, P. M., Zeltzer, L., & Harper, R. M. (2018).
774 A description of externally recorded womb sounds in human subjects during

- 775 gestation. *PloS one*, 13(5), e0197045.
776 <https://doi.org/10.1371/journal.pone.0197045>
- 777 Partanen, L., Korkalainen, N., Mäkikallio, K., Olsén, P., Laukkanen, P., &
778 Yliherva, A. (2018). Foetal growth restriction is associated with poor reading
779 and spelling skills at eight years to 10 years of age. *Acta Paediatrica*, 107(1), 79-
780 85. <https://doi.org/10.1111/apa.14005>
- 781 Querleu, D., Renard, X., Versyp, F., Paris-Delrue, L., & Crèpin, G. (1988). Fetal
782 hearing. *European Journal of Obstetrics & Gynecology and Reproductive*
783 *Biology*, 28(3), 191-212. [https://doi.org/10.1016/0028-2243\(88\)90030-5](https://doi.org/10.1016/0028-2243(88)90030-5)
- 784 Ribas-Prats, T., Almeida, L., Costa-Faidella, J., Plana, M., Corral, M. J., Gómez-Roig,
785 M. D., & Escera, C. (2019). The frequency-following response (FFR) to speech
786 stimuli: A normative dataset in healthy newborns. *Hearing Research*, 371, 28-
787 39. <https://doi.org/10.1016/j.heares.2018.11.001>
- 788 Ribas-Prats, T., Arenillas-Alcón, S., Lip-Sosa, D. L., Costa-Faidella, J., Mazarico,
789 E., Gómez-Roig, M. D., & Escera, C. (2022). Deficient neural encoding of
790 speech sounds in term neonates born after fetal growth restriction.
791 *Developmental Science*, 25(3), e13189. <https://doi.org/10.1111/desc.13189>
- 792 Ribas-Prats, T., Arenillas-Alcón, S., Pérez-Cruz, M., Costa-Faidella, J., Gómez-Roig,
793 M. D., & Escera, C. (2023a). Speech-encoding deficits in neonates born large-
794 for-gestational age as revealed with the envelope frequency-following response.

795 *Ear and hearing* 44(4), 829–841.

796 <https://doi.org/10.1097/AUD.0000000000001330>

797 Ribas-Prats, T., Cordero, G., Lip-Sosa, D. L., Arenillas-Alcón, S., Costa-Faidella, J.,
798 Gómez-Roig, M. D., & Escera, C. (in press). Developmental trajectory of the
799 frequency-following response during the first six months of life.

800 Rivera-Gaxiola, M., Silva-Pereyra, J., & Kuhl, P. K. (2005). Brain potentials to native
801 and non-native speech contrasts in 7- and 11-month-old American infants.
802 *Developmental science*, 8(2), 162-172. [https://doi.org/10.1111/j.1467-](https://doi.org/10.1111/j.1467-7687.2005.00403.x)
803 [7687.2005.00403.x](https://doi.org/10.1111/j.1467-7687.2005.00403.x)

804 Ruben, R. J. (1995). The ontogeny of human hearing. *International journal of pediatric*
805 *otorhinolaryngology*, 32, S199-S204. [https://doi.org/10.1016/0165-](https://doi.org/10.1016/0165-5876(94)01159-U)
806 [5876\(94\)01159-U](https://doi.org/10.1016/0165-5876(94)01159-U)

807 Russo, N. M., Skoe, E., Trommer, B., Nicol, T., Zecker, S., Bradlow, A., & Kraus, N.
808 (2008). Deficient brainstem encoding of pitch in children with autism spectrum
809 disorders. *Clinical Neurophysiology: official journal of the International*
810 *Federation of Clinical Neurophysiology*, 119(8), 1720-1731.
811 <https://doi.org/10.1016/j.clinph.2008.01.108>

812 Schneider, U., Schleussner, E., Haueisen, J., Nowak, H., & Seewald, H. J. (2001).
813 Signal analysis of auditory evoked cortical fields in fetal

- 814 magnetoencephalography. *Brain topography*, 14, 69-80.
815 <https://doi.org/10.1023/A:1012519923583>
- 816 Sharma, M., Bist, S. S., & Kumar, S. (2016). Age-related maturation of wave V latency
817 of auditory brainstem response in children. *Journal of Audiology & Otology*,
818 20(2), 97. <https://doi.org/10.7874/jao.2016.20.2.97>
- 819 Stuart, A., Yang, E. Y., & Green, W. B. (1994). Neonatal auditory brainstem response
820 thresholds to air-and bone-conducted clicks: 0 to 96 hours postpartum. *Journal*
821 *of the American Academy of Audiology*, 5(3), 163-172.
- 822 Su, P., Kuan, C. C., Kaga, K., Sano, M., & Mima, K. (2008). Myelination progression
823 in language-correlated regions in brain of normal children determined by
824 quantitative MRI assessment. *International Journal of Pediatric*
825 *Otorhinolaryngology*, 72(12), 1751–1763.
826 <https://doi.org/10.1016/j.ijporl.2008.05.017>
- 827 The MathWorks Inc. (2019). MATLAB version: 9.7 (R2019b), Natick, Massachusetts:
828 The MathWorks Inc. <https://www.mathworks.com>
- 829 Tsao, F. M., Liu, H. M., & Kuhl, P. K. (2004). Speech perception in infancy predicts
830 language development in the second year of life: A longitudinal study. *Child*
831 *development*, 75(4), 1067-1084. [https://doi.org/10.1111/j.1467-](https://doi.org/10.1111/j.1467-8624.2004.00726.x)
832 [8624.2004.00726.x](https://doi.org/10.1111/j.1467-8624.2004.00726.x)

- 833 Tsao, F. M., Liu, H. M., & Kuhl, P. K. (2006). Perception of native and non-native
834 affricate-fricative contrasts: Cross-language tests on adults and infants. *The*
835 *Journal of the Acoustical Society of America*, 120(4), 2285-2294.
836 <https://doi.org/10.1121/1.2338290>
- 837 Van Dyke, K. B., Lieberman, R., Presacco, A., & Anderson, S. (2017). Development of
838 phase locking and frequency representation in the infant frequency-following
839 response. *Journal of Speech, Language, and Hearing Research*, 60(9), 2740-
840 2751. https://doi.org/10.1044/2017_JSLHR-H-16-0263
- 841 Voegtline, K. M., Costigan, K. A., Pater, H. A., & DiPietro, J. A. (2013). Near-term
842 fetal response to maternal spoken voice. *Infant Behavior and Development*,
843 36(4), 526-533. <https://doi.org/10.1016/j.infbeh.2013.05.002>
- 844 Webb, A. R., Heller, H. T., Benson, C. B., & Lahav, A. (2015). Mother's voice and
845 heartbeat sounds elicit auditory plasticity in the human brain before full
846 gestation. *Proceedings of the National Academy of Sciences*, 112(10), 3152-
847 3157. <https://doi.org/10.1073/pnas.1414924112>
- 848 Werker, J. F., Gilbert, J. H., Humphrey, K., & Tees, R. C. (1981). Developmental
849 aspects of cross-language speech perception. *Child development*, 349-355.
850 <https://doi.org/10.2307/1129249>

- 851 Werker, J. F., & Tees, R. C. (1984). Cross-language speech perception: Evidence for
852 perceptual reorganization during the first year of life. *Infant behavior and*
853 *development*, 7(1), 49-63. [https://doi.org/10.1016/S0163-6383\(84\)80022-3](https://doi.org/10.1016/S0163-6383(84)80022-3)
- 854 White-Schwoch, T., Davies, E. C., Thompson, E. C., Carr, K. W., Nicol, T., Bradlow,
855 A. R., & Kraus, N. (2015b). Auditory-neurophysiological responses to speech
856 during early childhood: Effects of background noise. *Hearing research*, 328, 34-
857 47. <https://doi.org/10.1016/j.heares.2015.06.009>
- 858 White-Schwoch, T., Woodruff Carr, K., Thompson, E. C., Anderson, S., Nicol, T.,
859 Bradlow, A. R., Zecker, S. G., & Kraus, N. (2015a). Auditory processing in
860 noise: A preschool biomarker for literacy. *PLoS biology*, 13(7), e1002196.
861 <https://doi.org/10.1371/journal.pbio.1002196>
- 862 Wong, P. C., Skoe, E., Russo, N. M., Dees, T., & Kraus, N. (2007). Musical experience
863 shapes human brainstem encoding of linguistic pitch patterns. *Nature*
864 *neuroscience*, 10(4), 420-422. <https://doi.org/10.1038/nn1872>
- 865 Wyper, K. R., & Rasmussen, C. R. (2011). Language impairments in children with fetal
866 alcohol spectrum disorder. *Journal of Population Therapeutics and Clinical*
867 *Pharmacology*, 18(2). Retrieved from
868 <https://www.jptcp.com/index.php/jptcp/article/view/485>
- 869