The frequency-following response (FFR) to speech stimuli: a normative dataset in healthy newborns

Teresa Ribas-Prats[1,2,3], Laura Almeida[3,4], Jordi Costa-Faidella[1,2,3], Montse Plana[3,4], MJ Corral[1,2,3], M. Dolores Gómez-Roig[3,4], Carles Escera[1,2,3]

[1] Brainlab – *Cognitive Neuroscience Research Group.* Department of Clinical Psychology and Psychobiology, University of Barcelona (Catalonia-Spain)

[2] Institute of Neurosciences, University of Barcelona (Catalonia-Spain)

[3] Institut de Recerca Sant Joan de Déu, Esplugues de Llobregat (Barcelona, Catalonia-Spain)

[4] BCNatal – Barcelona Center for Maternal Fetal and Neonatal Medicine (Hospital Sant Joan de Déu and Hospital Clínic), University of Barcelona, Barcelona (Catalonia-Spain).

Abbreviated title: Normative FFR in neonates

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

Number of pages: 22; Number of words in Abstract: 197; Number of words in the manuscript: 8214. Number of figures: 8. Number of tables: 2.

Declarations of interest: none.

Highlights

- Frequency-following responses (FFR) were recorded on 46 newborns
- Seven objective FFR parameters were retrieved in time and frequency domains
- A normative data-base is offered to guide future clinical studies

Abstract

The Frequency-Following Response (FFR) is a neurophonic auditory evoked potential that reflects the efficient encoding of speech sounds and is disrupted in a range of speech and language disorders. This raises the possibility to use it as a potential biomarker for literacy impairment. However, reference values for comparison with the normal population are not yet established. The present study pursues the collection of a normative database depicting the standard variability of the newborn FFR. FFRs were recorded to /da/ and /ga/ syllables in 46 neonates born at term. Seven parameters were retrieved in the time and frequency domains, and analyzed for normality and differences between stimuli. A comprehensive normative database of the newborn FFR is offered, with most parameters showing normal distributions and similar robust responses for /da/ and /ga/ stimuli. This is the first normative database of the FFR to characterize normal speech sound processing during the immediate postnatal days, and corroborates the possibility to record the FFRs in neonates at the maternity hospital room. This normative database constitutes the first step towards the detection of early FFR abnormalities in newborns that would announce later language impairment, allowing early preventive measures from the first days of life.

Keywords: auditory brainstem response, FFR, auditory processing, speech encoding, hearing screening, human neonates, language impairments, cognition.

1. Introduction

Auditory brainstem responses (ABR) otoacoustic emissions (OAE) to simple auditory stimuli are used in clinical practice to assess auditory pathway integrity (Hoth et al., 2009; Moller, 1999; Stuart et al., 1996). Universal newborn hearing screening based on the ABR or OAE has provided the earliest possible diagnosis for infants with permanent hearing loss (Moeller et al., 2006). ABRs can also be elicited with more complex stimuli such as speech or music. Studying brainstem responses to complex stimuli has revealed that the subcortical auditory system is more than a set of relay stations transmitting information from the peripheral sensory epithelium to the cerebral cortex (Song et al., 2008). Subcortical contributions to the neural representation of sound can hence be reliably studied using a variant of the ABR, the so-called Frequency-Following Response (FFR).

The FFR is a non-invasive electrophysiological measurement that reflects the encoding of the temporal and spectral characteristics of complex evoking sounds in a cortico-subcortical auditory network (Bidelman, 2018; Kraus and White-Schwoch, 2015; Skoe and Kraus, 2010). Disruptions in the FFR are found in children with deficits in phonological awareness, reading and abnormal timing resolution (Abrams et al., 2006; Banai et al., 2009, 2005; Basu et al., 2010; Hornickel et al., 2012, 2011, 2009; Johnson et al., 2007; King et al., 2002; Kraus et al., 1996; Wible et al., 2004). Children with reading or language disorders have significantly slower neural response timing, weak neural encoding of formant-related stimulus harmonics and less robust tracking of frequency contours than typically developing children (Banai et al., 2005; Basu et al., 2010; Billiet and Bellis, 2010; Hornickel et al., 2009; Wible et al., 2004).

In addition, neurodevelopmental disorders characterized by impaired communication and literacy skills such as dyslexia or autism spectrum disorder (ASD) have been associated with abnormal subcortical representation of speech sounds. Hornickel and Kraus (2013) described that children with dyslexia are characterized by delayed and harmonically impoverished responses from the auditory brainstem (Banai et al., 2009; Basu et al., 2010; Hornickel et al., 2012, 2011) and reduced subcortical representation of stimulus differences (Banai et al., 2005; Hornickel et al., 2009). In children with ASD, Russo et al. (2009) reported deficits in timing and frequency encoding of speech sounds at brainstem level. ASD individuals also exhibited a subcortical neural response which was more vulnerable to background noise in comparison to typically developing children. Another recent study described more variable FFRs in children with ASD compared to their healthy peers (Otto-Meyer et al., 2018).

Moreover, White-Schwoch and colleagues (2015a) examined the FFRs recorded from a group of 3-4 years old children and reported that the precision and stability of the neural encoding of consonants in noise predicted performance on reading readiness tests one year after the neurophysiological assessment. In infants with ages between 3 to 10 months, Anderson et al. (2015) found a robust fundamental frequency representation while the amplitudes of the first formant and higher harmonics in the FFR increased with age. Thus, the FFR may offer a neurophysiological marker of the efficient encoding of speech sounds related to literacy abilities (Hornickel and Krauss, 2013).

In spite of the potential of the FFR to anticipate reading and literacy impairments, little is known about the neural transcription of speech sounds in newborns. Recently, Kraus and

White-Schwoch (2016) suggested an idealized scenario where the FFR could be registered after the hearing screening test was passed using the same equipment, with only replacing the typical click or burst stimulus used for ABR with a speech sound. Their claim was that FFR screening could help identifying newborns at risk of developing literacy impairments in the future, so that they could be referred to appropriate specialists for a more exhaustive examination.

The interest to improve our knowledge in the early maturation of complex sound processing at brainstem level is further encouraged by the plasticity of neural tissues in subcortical structures (Johnson et al., 2008; Krisham et al., 2009, 2005; Musacchia et al., 2007; Russo et al., 2005; Song et al., 2008; Strait et al., 2009; Wong et al., 2007). Recent studies have shown that the FFR is sensitive to language (Krishnan et al., 2009, 2005; Krizman, 2015; Xu et al., 2006), musical experience (Bidelman et al., 2014; Jonhson et al., 2008; Lee et al., 2009; Musacchia et al., 2007; Strait et al., 2009; Weiss and Bidelman, 2015; Wong et al., 2007) and short- and long-term auditory adaptation and training (Escera, 2017; Russo et al., 2005; Song et al., 2008). Evidence of brainstem response modulation by musical training has been shown with a few years of lessons (Parbery-Clark et al., 2012; Strait and Kraus, 2014; Zuk et al., 2013) and observed in very young musicians of 3 years of age (Skoe and Kraus, 2013; Strait et al., 2014, 2013, 2012). After a revision of a number of phonologic intervention studies, Handler et al. (2011) confirmed the effectiveness of phonologic training implementation during the first years of life in reading impaired children as compared to those who were not identified or helped until years later (Foorman et al., 2003; Lyon et al., 2010; National Institutes of Health, 2000; Schatschneider and Torgesen, 2004; Vellutino et al., 2004; Willcutt and Pennington, 2000). Hence, detecting early FFR abnormalities in newborns that would announce later language impairment would allow installing early preventive measures such as music enrichment programs (Kraus et al., 2014).

The possibility to record FFR in newborns has been suggested in no more than ten studies so far, showing a remarkable similarity between the adult and newborn responses (Gardi et al., 1979; Jeng et al., 2016, 2013, 2011b). In addition, Jeng et al. (2016, 2013, 2011b) reported similar FFRs in American and Chinese newborns indicating an early and universal maturation of voice-pitch processing at brainstem level in neonates at 1-3 days after birth. However, before the FFR can be used in the clinics, normative values need to be established. The aim of the present study was the collection of a normative database depicting the standard variability observed in different relevant parameters retrieved from the newborn FFR.

2. Methods

2.1. Participants

Fifty term newborns (24 females, aged 14-125 hours after birth) were recruited from *Sant Joan de Déu Hospital* in Barcelona (Catalonia, Spain). By medical reports, all newborns were low-risk gestations without neither obstetric pathologies nor risks factors for hearing impairment as determined by the Committee on Infant Hearing as a “high-risk registry” (Joint Committee on Infant Hearing, 1994). All the participants had a high Apgar score (>7) at 1 and 5 minutes of life.

All newborns passed the standardized hearing screening of peripheral auditory health using an automated auditory brainstem response system (ALGO 3i, Natus Medical Incorporated, San

Carlos, CA), before the experiment. Each newborn also passed a normal click-evoked auditory brainstem response collected with a standard SmartEP platform (Intelligent Hearing Systems, Miami, Fl). Standard click ABR was recorded in response to a 100 µs square-wave click stimulus presented at 55 dB SPL according to standard procedures to the Joint Committee on Infant Hearing (2007). Mean latency and amplitude of wave V were 8.90 (± 0.79 SD) ms and 0.27 (± 0.15 SD) µV, respectively (Fig. 1), being comparable to the norms published by Stuart et al. (1994). Four newborns were excluded from the final sample because their wave V could not be reliably identified.

Informed consent was obtained from legal guardians of the newborns assessed in accordance with the Declaration of Helsinki after the study has approved by the Ethical Committee of Clinical Research (CEIC) of the Sant Joan de Déu Foundation.

2.2. Stimuli

The complex auditory stimuli used were the syllables /da/ and /ga/ created with a Klatt-based synthesizer (Klatt, 1976) and modified by Praat (Boersma and Weenink, 2013) to have a fundamental frequency (F_0) of 113 Hz in both syllables. The use of F_0 of 113 Hz instead of the very common F_0 at 100 Hz was motivated to avoid contamination by harmonics of the 50 Hz electric line in Europe. The stimuli had a duration of 170 ms, including a 10 ms onset period, a 47 ms consonant transition and a 113 ms steady state vowel. These sounds differed only by the trajectory of their second formant (F2) during the consonant transition (/da/: 1438-1214 Hz; /ga/: 1801-1214 Hz). In the vowel region, all the formants were stable in both syllables. Stimuli were presented at 55 dB SPL in alternating polarities with a 100.27 ms inter-stimulus interval (presentation rate of 3.7 Hz) to the right ear through Etymotic shielded earphones of 300 ohms (ER, Elk Grove Village, IL, EEUU) connected to Flexicoupler® (Natus Medical Incorporated, San Carlos, CA) adaptor.

2.3. Procedure

All newborns were tested in the hospital room where they were resting with their mother once the newborn hearing screening was already passed. For the recording, each newborn was asleep in its own bassinet and the researcher interrupted the experiment to any sign of sleep disruption, to restart the recording as soon as the newborn was calm again. The total mean duration of a test session was 26.09 (± 3.56 SD) min (51.81 ms x 2000 sweeps x 1 condition + 270.27 ms x 2000 sweeps x 2 conditions + duration of sweeps rejected + time needed to check electrodes impedances + time for the newborn restored the sleep in case that it was disrupted). If the recording time exceeded this duration, the session was canceled and postponed for several hours or until the next day.

2.4. Data adquisition

FFRs were collected with a SmartEP platform including the *cABR* and *Advanced Hearing Research* modules (Intelligent Hearing Systems, Miami, Fl, EEUU). Auditory brainstem responses to click and to the syllables /da/ and /ga/ were recorded using disposable snap Ag/AgCl electrodes placed in a vertical montage (the active electrode was located at Fpz, references at each ear, and ground at the forehead) with impedances < 5 kΩ (Fig. 2). Only the ipsilateral reference was used in the analysis (Hornickel et al., 2009). The continuous EEG was online bandpass filtered

from 30-3000 Hz and acquired with a sampling rate of 13333 Hz. Responses were collected by alternating polarity and averaged ((Org.+Inv.)/2) to isolate the neural response by minimizing stimulus artifact and cochlear microphonic (Aiken and Picton, 2008). Any activity exceeding ±30 μV was rejected online as an artifact, and a total of 2000 artifact-free responses were averaged for each syllable and newborn. Rejection was smaller than 15% of sweeps per condition.

2.5. Data processing and analyses

Neural responses to /da/ and /ga/ were averaged separately and epoched into 270.27 ms windows (including 40 ms pre-stimulus period). After averaging, data were filtered off-line with a spectral bandpass filter with infinite slope from 80 to 3000 Hz to isolate brainstem responses (Chandrasekaran and Kraus, 2009).

After a review of the analytic techniques used to characterize the FFR in different populations (Anderson et al., 2015; Banai et al., 2009; Basu et al., 2010; Hornickel et al., 2012, 2011, 2009; Hornickel and Kraus, 2013; Jeng et al., 2013, 2011a, 2011b, 2010; Jonhson et al., 2008, 2007; Liu et al., 2015; Musacchia et al. 2007; Neef et al., 2017; Russo et al., 2005; 2004; Skoe et al., 2011; Strait et al., 2014; White-Schwoch et al., 2015a and 2015b; White-Schwoch and Kraus, 2013), an extensive number of parameters are included in this study to offer an accurate description of FFR properties in newborns. In the time domain, cross-correlation between the stimulus and the neural response, the neural lag and the signal-to-noise ratio (SNR) are reported. Pitch error and pitch strength, extracted from a sliding time-window autocorrelation approach, were also evaluated. In the spectral domain, the amplitude of the fundamental frequency and its harmonics was calculated based on fixed time-windows. Signal-to-noise ratio variation along the stimulus duration, computed as points below noise floor using a sliding time-window approach, is also reported. All analyses were performed under Matlab R2015b (Mathworks) using routines provided by Intelligent Hearing Systems (Miami, Fl, EEUU) and custom scripts developed in our laboratory. The following sections describe the rationale and procedure to obtain each of these measures.

2.5.1. Time domain

To examine the FFR in the time domain, a cross-correlation between each syllable stimulus and the neural response was computed. For the purpose of comparison, each syllable was resampled to the sampling rate of the EEG recording (13333 Hz), and bandpass filtered between 80 – 3000 Hz. The magnitude of the first maximum cross-correlation peak and its lag are reported. In addition, the signal-to-noise ratio (SNR) of the FFR was computed in three time windows, corresponding to three different regions of the acoustic stimulus: consonant transition (10 – 57 ms), vowel (57 – 170 ms) and entire response (0 – 170 ms). In the following lines, each of these objective parameters is briefly described.

2.5.1.1. Stimulus-to-response cross-correlation. The cross-correlation magnitude shows how faithfully the FFR reproduces the stimulus waveform as a function of the time shift between the two (Russo et al., 2004). The maximum cross-correlation value was obtained within a response time lag of 3 to 10 ms, in line with previous studies (Jeng et al., 2010).

2.5.1.2. Neural lag. The neural lag is an estimation of the transmission delay between stimulus and response. It was retrieved as the time lag that produced the maximum stimulus-to-response cross-correlation magnitude computed as described above.

2.5.1.3. Signal-to-noise ratio (SNR). Root mean square (RMS) amplitude (in μV) indicates the overall magnitude of neural activity over time (Liu et al., 2015). RMS is calculated by squaring each point in a region of the response waveform, computing the mean of the squared values and then computing its square root. The RMS amplitudes of the FFR to the consonant transition, the vowel and the entire stimulus were computed separately and divided by the RMS amplitude of the pre-stimulus period (-40 – 0 ms) (Anderson et al., 2015; Russo et al., 2004). In order to account for individual response-to-stimulus delays, the individual neural lag retrieved from the stimulus-response cross-correlation analysis was added to each of the selected regions (except to the pre-stimulus region).

Additionally, we computed the autocorrelation function of the entire FFR recording in a sliding time-window approach, using short bins of 40 ms (Hanning tapered) with 1 ms step size. In each bin, the periodicity of the F_0 was estimated as the frequency (inverse of the lag) yielding the maximum autocorrelation value (Boersma, 1993) within a predefined frequency range, fixed from 103 to 123 Hz, leaving a 10 Hz buffer at each side of the F_0 frequency of the stimulus (113 Hz) (Jeng et al., 2013, 2011b, 2010). The F_0s from each bin were concatenated to construct the F_0 contour of the FFR recording. The same procedure was applied to the stimulus waveforms /da/ and /ga/. Two objective parameters were extracted using this method, as described next.

2.5.1.4. Pitch error. Pitch error (in Hz) is a measure of pitch encoding accuracy of the FFR along the syllable presentation. Once the syllable was shifted the time equivalent to the average of the individual neural lag, the absolute Euclidian distance between the syllable F_0 and the response F_0 from each bin was calculated and averaged together (Liu et al., 2015; Song et al., 2008).

2.5.1.5. Pitch strength. Pitch strength is a measure of periodicity that reflects the robustness of the response's phase-locking to the syllable F_0 contour (Jeng et al, 2013). It was computed as the average across bins of the normalized autocorrelation values at the signal's F_0.

2.5.2. Frequency domain

To examine the FFR in the frequency domain, we computed a Fast Fourier Transform (FFT; Hamming windowed) separately for the three different regions of the stimuli as described above (*see 2.5.1*), each region adjusted for each newborn to account for the individual neural lag. In order to avoid artifact differences in power estimates due to different window lengths, the Welch's averaging method (Welch, 1967; *pwelch.m* Matlab function) was applied to estimate the power spectrum of the demeaned signal corresponding to the consonant transition, the vowel and the entire response (segments of 40 ms, equivalent to the window length of the pre-stimulus region; Hamming windowed; 82.5% overlap; spectrum type specified as 'power').

2.5.2.1. Spectral amplitude. The spectral amplitude indicates the magnitude of neural phaselocking at a certain frequency (White-Schwoch et al., 2015b). In order to compare our results with the previous literature, a square root of power estimates was applied to convert them to spectral amplitude. Spectral amplitude corresponding to the F_0 (113 Hz) and its integer

harmonics up to 1500 Hz (i.e., HH_{2-13}) were computed as the mean over a ± 5 Hz frequency window centered at each individual peak. Two values are reported: 1) the F_0 spectral amplitude; and 2) a composite value for all F_0 harmonics, resulting from averaging the spectral amplitudes of HH_{2-13} (Krizman et al., 2015; Parbery-Clark et al., 2009).

2.5.2.2. Points below noise floor. In addition, we computed the spectrogram of the entire FFR recording in short bins of 40 ms (Hanning tapered) with 1 ms step size. Spectral amplitudes were computed with FFTs after zero-padding each FFR bin to 1s to increase spectral resolution, and Hanning windowed. This method allowed extracting the points below noise floor, which is an objective index that describes how the signal could be discerned and differentiated from the noise. It is computed dividing the F_0 spectral amplitude of each bin by the F_0 spectral amplitude computed in the pre-stimulus region (Song et al., 2008). The total number of bins in which the spectral amplitude of F_0 is smaller in the post- vs. the pre-stimulus region is reported.

2.6. Statistical analyses

SPSS 22.0 was used for statistical analysis. Descriptive statistics for each parameter computed for each syllable (/da/, /ga/) are presented as mean and standard deviation (SD) or median and interquartile range (IQR), after assessing normality distribution by Kolmogorov-Smirnov Statistic with Lilliefors' Significance.

The initial statistical approach was a 2x2 repeated measures analysis of variance (RMANOVA), with syllable (/da/, /ga/) and region (consonant transition, vowel) as factors. However, in the objective indices where the results were obtained only in one response region (i.e., stimulus-to-response cross-correlation, neural lag, pre-stimulus RMS amplitude, pitch error, pitch strength and points below noise floor), a t-test or Wilcoxon test was computed according to the assumption of normal distribution. In the spectral amplitude analyses, a third factor was included (harmonic: F_0, average amplitude HH_{2-13}). The Greenhouse–Geisser correction was applied when the assumption of sphericity was violated. A result was considered significant when $p<0.05$ using a two-tailed analysis. Bonferroni correction was used to adjust p-values for all multiple pairwise contrasts.

3. Results

Grand-average FFR waveforms for both syllables (/da/, /ga/) and the corresponding amplitude spectra of the consonant transition and the vowel regions of the response are shown in Fig. 3A. Individual waveforms and corresponding amplitude spectra from a typical newborn with high SNR (NF 029) and from another newborn with very low SNR (NF 038) are plotted in Fig. 3B and 3C, respectively. Grand-average spectrograms are presented in Fig. 4A, and those from the same individual newborns as in Fig. 3 are shown in Fig. 4B and 4C. As Fig. 4 shows, the newborn FFR contains clear energetic content in the F_0 contour of the syllable, but not in its harmonic frequencies, in agreement with Jeng et al. (2016, 2013, 2011b). Below, we provide a series of descriptive statistics extracted from the values obtained to all computed parameters. In addition, we provide a statistical comparison of parameter values across stimulus types (syllable /da/ vs. /ga/).

3.1. Time domain

In the time domain, a normal distribution was observed for most of the parameters assessed. Only the values of the SNR computed for the consonant transition when the syllable /da/ was presented and the values of the SNR calculated for the vowel and the entire response when the syllable /ga/ was presented did not follow a normal distribution. The corresponding descriptive statistics are reported in Table 1. Fig. 5 illustrates the distribution of the values obtained in all time-domain parameters for each syllable (/da/, /ga/) and indicates the response region assessed (consonant transition and vowel) for SNR. Pitch tracking and pitch strength obtained from the grand-average waveforms and from the individual newborns depicted in Fig. 3 with high and low SNR are shown in Fig. 6 and 7, respectively.

A comparison of the time-domain parameter values obtained to the /da/ and /ga/ syllables showed no significance differences in stimulus-to-response cross-correlation values ($t(45) = 0.032$, $p = 0.974$) nor in neural lag ($t(45) = -0.251$, $p = 0.803$). RMS amplitudes during the pre-stimulus baseline period were not significantly different either ($t(45) = 1.860$, $p = 0.069$). Thus, any difference observed in response to the /da/ and /ga/ stimuli could not be attributed to variation in pre-stimulus spontaneous activity. However, no differences were observed in SNR between syllables ($F(1,45) = 1.630$, $p = 0.208$, $\eta^2 = 0.239$). Regarding response region, in contrast with previous studies, the SNR of the vowel was smaller than that of the consonant transition, irrespective of the syllable ($F(1,45) = 32.385$, $p < 0.001$, $\eta^2 = 1$). Finally, pitch measurements, extracted from a sliding time-window autocorrelation approach applied to the entire FFR recording, did not show any significant differences between syllables either (pitch error: $t(45) = -1.440$, $p = 0.157$; pitch strength: $t(45) = -1.187$, $p = 0.242$).

3.2. Frequency domain

In the frequency domain, spectral amplitude values followed a normal distribution for most of the conditions assessed. However, points below noise floor values did not pass the Kolmogorov-Smirnov test. Descriptive statistics are reported in Table 2. Fig. 8 illustrates the distribution of spectral amplitude values across response regions (consonant transition, vowel), harmonics (F_0, $HH_{2\text{-}13}$) and syllables (/da/, /ga/), as well as the distribution of points below noise floor values across syllables. Fig. 6 shows the points below noise floor computed from the grand-average waveforms and from the individual newborns depicted in Fig. 3 with high and low SNR.

A three-factor RMANOVA (syllable: /da/, /ga/; region: consonant transition, vowel; harmonic: F_0, $HH_{2\text{-}13}$) performed on all spectral amplitude values showed no significance differences between syllables ($F(1,45) = 0.038$, $p = 0.846$, $\eta^2 = 0.054$). Regarding response region, spectral amplitudes were larger to the consonant transition than to the vowel ($F(1,45) = 22.008$, $p < 0.001$, $\eta^2 = 0.996$) in agreement with previous studies (White-Schwoch et al., 2015b). Regarding the harmonic factor, an expected significant effect was found in spectral amplitude ($F(1,45) = 274.123$, $p < 0.001$, $\eta^2 = 1$), being larger for the F_0 than for its harmonics. A significant interaction between the region and the harmonic was observed as well ($F(1,45) = 23.631$, $p < 0.001$, $\eta^2 = 0.997$). Post-hoc analyses showed a slightly higher difference in spectral amplitudes between the F_0 and those of the $HH_{2\text{-}13}$ in the consonant transition (mean F_0 = 43.293 nV, SE = 2.801; mean $HH_{2\text{-}13}$ = 2.892 nV, SE = 0.110) than in the vowel portion (mean F_0 = 35.212 nV, SE = 1.902; mean $HH_{2\text{-}13}$ = 2.917 nV, SE = 0.101). Finally, no significant differences were observed in points below noise floor across syllables ($Z = -1.116$, $p = 0.264$).

4. Discussion

We here provide the first normative database of the newborn FFR, thus demonstrating the feasibility to record this electrophysiological response from bedside, in the maternity unit where newborns are delivered, during their first hours of life.

In order to obtain a comprehensive database, two syllables with different stop consonants (/da/, /ga/) were presented to 50 newborns aged 14-125 hours. Descriptive statistics of each of the parameters assessed were obtained (see Table 1-2). Normal distribution was confirmed for most of them, for which mean and standard deviation are reported. In those cases in which parameter values did not follow a normal distribution, median and interquartile range are reported instead. The representation of the distribution of all values using quartiles opens the possibility to determine the location of an individual's scores and to assess if her/his results are included into the central 50% of values. In general, comparing our results with previous literature, we found smaller values in all measures (Skoe et al., 2013; Strait et al., 2013; White-Schwoch et al., 2015b) except in pitch error and pitch strength (Jeng et al., 2010). The lower values obtained, according to Skoe et al. (2013) could be attributed to age differences.

For each of the objective parameters assessed, no differences between /da/ and /ga/ syllables were observed, indicating that these two stimuli could be used interchangeably to obtain a potential snapshot of pitch representation in the newborn's brain before they are discharged. However, differences were found between parameter values extracted from distinct portions of the FFR, which corresponded to different stimulus regions, and from different harmonic components of the frequency spectrum. For instance, higher SNR and spectral amplitude values were found in the consonant transition compared to the vowel, possibly due to response adaptation. In contrast, White-Schwoch et al. (2015b) showed higher SNR values for the vowel region than for the consonant transition, although comparable results in spectral amplitude. Regarding the harmonic content, higher spectral amplitude values were observed for the F_0 in comparison to higher harmonics (see Fig. 8), in line with previous observations (Anderson et al., 2015; Banai et al., 2009; Jeng et al., 2016, 2013, 2011b; White-Schwoch et al., 2015b). While we cannot rule out that this may be a particular characteristic of the newborn FFR, the procedure used to obtain it, by averaging responses to opposite polarities, may have exerted a major influence. In fact, while this operation enhances envelope representation, which is not polarity dependent, it eliminates the temporal fine structure of the signal, drastically reducing the spectral energy of the present harmonics (Aiken and Picton, 2008). This explanation may account as well for the lack of differences found in parameter values across syllables.

The results obtained here are in line with those of Jeng and colleagues (2016, 2013, 2011b) who corroborated that the newborn FFR (age: 1-3 days) could be registered during the first days of life (Gardi et al., 1979). In addition to finding evident spectral energy content at the F_0 in the newborn FFR spectrograms, differences in the F_0 spectral energetic content across the individual newborns assessed were observed (Jeng at al., 2016, 2013, 2011b), as in our study. However, descriptive statistics of the different FFR parameters as the ones computed in the present study were only published in Jeng et al. (2010), specifically for measurements related to pitch tracking. They found similar FFRs to voice pitch between infants of 5.7 months mean age and adults using a Chinese monosyllabic stimulus that mimics the English vowel /i/ with a rising pitch (117 to 166

Hz). Mean and standard deviation of pitch error and pitch strength recorded from the infant group were slightly smaller than those found in the present study. In spite of the differences in group age and the stimulus materials, our results, together with those of Jeng et al. (2010), confirm that pitch is accurately processed during the first stages of life.

We believe our study contributes a step forward towards achieving the ideal scenario proposed by Kraus and White-Schwoch (2016). Taking the universal newborn hearing screening (UNHS) as a reference, they proposed the implementation of a cognitive screening after the UNHS is passed, by using the same recording system where only the classical click stimulus would be replaced by a speech sound. In this study, FFRs were recorded in the hospital room where newborns were resting with their mothers after the auditory screening was passed. Auditory pathway integrity was corroborated by Wave V identification to a click presentation (Fig. 1) using the same portable equipment employed afterwards to record the FFR to the syllables /da/ and /ga/. As the specific hearing screening protocol recommended by the Joint Committee of Infant hearing (2007) to promote the early identification and intervention of children with hearing loss, we suggest, according to Kraus and White-Schwoch (2016), the establishment of a protocol with which the identification of a disrupted newborn FFR would prompt the child to be referred to a coordinated team for follow up, in seeking intervention strategies to improve his/her literacy abilities. Several studies support that spectro-temporal encoding mechanisms could be improved after short- (Song et al., 2012) and long-term (Hornickel et al., 2012) training protocols using the FFR as a fingerprint reflecting these improvements. The neural plasticity of the subcortical auditory system during the first days of life highlights the importance to this period to detect as earlier as possible an abnormal FFR.

The speech tokens selected (/da/ and /ga/) are extensively used in the literature because several studies suggest that stop consonants are an important constraint in populations with literacy impairments (Kraus et al., 1996; Tallal and Stark, 1981; Turner et al., 1992). However, recent studies suggest that the identification of the speech in noise is especially compromised in populations with specific language impairment (Cunningham et al., 2001; Hornickel et al., 2009, White-Schwoch et al., 2015a, 2015b, 2013). To complement the present normative database, future studies using different stimulus materials, for example including speech in noise, and achieving other potential objective indices retrieved from the FFR are needed.

In any case, this study offers valuable information relative to a specific lifetime that was not included in previous normative database research and remained to be described. Skoe and colleagues (2013) characterized the auditory brainstem response of 586 healthy participants across an extensive age range (from 3 months to 72 years). With our data we contribute to fill the gap in the research focused on the FFR trends along the lifetime. In addition, an age-appropriate normative database is critical to establish a reference with which to compare the results of an individual from a population that presents risk factors to develop literacy impairments (Jeng et al., 2016; Skoe et al., 2013).

Albeit the potential clinical utility attributed to the FFR, its small amplitude at the scalp in contrast with the high background noise constraints the quality of the recordings to translate the use of the FFR from brain research to clinical applications. The averaging method is usually used in clinical practice to assess how the signal is discernible to the noise recorded as a function

of the number of sweeps presented (Jeng et al., 2011a). Recent studies offered different threshold criterion depending on the statistical approach used and population assessed (Bidelman et al., 2018, 2014; Jeng et al., 2018, 2013, 2011a). Thus, before the FFR could be considered a universal clinical test, a consistent predetermined threshold to verify a distinguishable FFR is required for each specific lifetime (Jeng et al., 2016).

In summary, the present study shows the possibility to record newborn FFR during the first postnatal hours at the maternity unit before discharge and contributes to approximate the FFR to clinical context. In agreement with Jeng et al. (2016), our study promotes longitudinal research in which the newborns that present normal and abnormal FFR were followed up to elucidate whether an FFR recorded during the first days of life could become a biomarker to prevent future literacy disorders.

Acknolwedgements.

The authors want to express their warmest thanks to all newborns that peacefully slept in their basinet during our recordings, and to their mothers who generously agreed to participate in the study. Also thanks go for Francisco Díaz-Santaella for inestimable technical support in developing the scripts for data analysis. This work was supported by the 2017SGR-974 Excellence Research Group of the Generalitat de Catalunya, the PSI2015-63664-P project (MINECO/FEDER), and the ICREA Acadèmia Distinguished Professorship awarded to Carles Escera.

References

Abrams, D.A., Nicol, T., Zecker, S.G., Kraus, N., 2006. Auditory brainstem timing predicts cerebral dominance for speech sounds. J Neurosci. 26, 11131–11137.

Aiken, S.J., Picton, T.W., 2008. Envelope and spectral frequency-following responses to vowel sounds. Hear Res. 245, 35–47.

American Academy of Pediatrics, Joint Committee on Infant Hearing, 2007. Year 2007 Position Statement: Principles and Guidelines for Early Hearing Detection and Intervention Programs. Pediatrics 120, 898 - 921.

Anderson, S., Parbery-Clark, A., White-Schwoch, T. Kraus, N., 2015. Development of subcortical speech representation in human infants. J Acoust Soc Am. 137, 3346–3355.

Banai, K., Hornickel, J., Skoe, E., Nicol, T., Zecker, S., Kraus, N., 2009. Reading and subcortical auditory function. Cereb Cortex. 19, 2699–2707.

Banai, K., Nicol, T., Zecker, S.G., Kraus, N., 2005. Brainstem Timing: Implications for Cortical Processing and Literacy. J Neurosci. 25, 9850–9857.

Basu, M., Krishnan, A., Weber-Fox, C., 2010. Brainstem correlates of temporal auditory processing in children with specific language impairment. Dev Sci. 13, 77–91.

Bidelman, G. M., 2018. Subcortical sources dominate the neuroelectric auditory frequency-following response to speech. NeuroImage 175, 56-69

Bidelman, G.M., Weiss, M.W., Moreno, S., Alain, C., 2014. Coordinated plasticity in brainstem and auditory cortex contributes to enhanced categorical speech perception in musicians. Eur J Neurosci. 40, 2662–2673.

Billiet, C.R., Bellis, T.J., 2010. The relationship between brainstem temporal processing and performance on tests of central auditory function in children with reading disorders. J Speech Lang Hear Res. 54, 228-242.

Boersma, P., 1993. Accurate short-term analysis of the fundamental frequency and the harmonics-to-noise ratio of a sampled sound. IFA Proceedings 17, 97–110.

Boersma, P., Weenink, D., 2013. Praat: doing phonetics by computer [Computer program]. Version 5.3.51, retrieved 2 June 2013 from http://www.praat.org/.

Chandrasekaran, B., Kraus, N., 2009. The scalp-recorded brainstem response to speech: neural origins and plasticity. Psychophysiology 47, 236-246.

Cunningham, J., Nicol, T., Zecker, S., Kraus, N., 2001. Neurobiologic responses to speech in noise in children with learning problems: deficits and strategies for improvement. Clin Neurophysiol. 112, 758-767.

Escera, C., 2017. The Role of the Auditory Brainstem in Regularity Encoding and Deviance Detection. In: Kraus N, Anderson S, White-Schwoch T, Fay R and Popper A, editors. The Frequency Following Response: A Window into Human Communication. Springer International Publishing.

Foorman, B.R, Breier, J.I., Fletcher, J.M., 2003. Interventions aimed at improving reading success: an evidence-based approach. Dev Neuropsychol. 24, 613– 639.

Gardi, J., Salamy, A., Mendelson, T., 1979. Scalp-recorded frequency-following responses in neonates. Audiology 18, 494–506.

Handler, S.M. Fierson W.M., Section on Ophthalmology; Council on Children with Disabilities; American Academy of Ophthalmology; American Association for Pediatric Ophthalmology and Strabismus; American Association of Certified Orthoptists., 2011. Learning disabilities, dyslexia, and vision. Pediatrics 127, 818 – 856.

Hornickel, J., Chandrasekaran, B., Zecker, S., Kraus, N., 2011. Auditory brainstem measures predict reading and speech-in-noise perception in schoolaged children. Behav Brain Res. 216, 597–605.

Hornickel, J., Kraus, N., 2013. Unstable representation of sound: a biological marker of dyslexia. J Neurosci. 33, 3500–3504.

Hornickel, J., Skoe, E., Nicol, T., Zecker, S., Kraus, N., 2009. Subcortical differentiation of stop consonants relates to reading and speech-in-noise perception. Proc Natl Acad Sci U S A. 106, 13022–13027.

Hornickel, J., Anderson, S., Skoe, E., Yi, H.G., Kraus, N., 2012. Subcortical representation of speech fine structure relates to reading ability. Neuroreport 23, 6–9.

Hoth, S., Neumann, K., Weisschuh, H., Bräunert, J., Böttcher, P., Hornberger, C., Maul, H., Beedgen, B., Buschmann, K., Sohn, C., Hoffmann, G., Plinkert, P. (2009). Universal newborn hearing screening. Methodical aspects. HNO, 57,29-36.

Jeng, F.C. Nance, B., Montgomery-Reagan, K., Lin C.D., 2018. Exponential modeling of frequency-following responses in American neonates and adults. J Am Acad Audiol. 29, 125-134.

Jeng, F.C., Hu, J., Dickman, B., Lin, C.Y., Lin, C.D., Wang, C.Y., Chung, H.K., Li, X., 2011a. Evaluation of two algorithms for detecting human frequency-following responses to voice pitch. Int J Audiol. 50, 14–26.

Jeng, F.C., Hu, J., Dickman, B.M., Montgomery-Reagan, K., Tong, M., Wu, G., Lin, C.D., 2011b. Cross-linguistic comparison of frequency-following responses to voice pitch in American and Chinese neonates and adults. Ear Hear. 32, 699–707.

Jeng, F.C., Lin, C.D., Wang, T.C., 2016. Subcortical neural representation to Mandarin pitch contours in American and Chinese newborns. J Acoust Soc Am. 139, 190-195.

Jeng, F.C., Peris, K.S., Hu, J., Lin, C.D., 2013. Evaluation of an automated procedure for detecting frequency-following responses in American and Chinese neonates. Percept Mot Skills. 116, 456–465.

Jeng, F.C., Schnabel, E.A., Dickman, B.M., Hu, J., Li, X., Lin, C.D., Chung, H.K., 2010. Early maturation of frequency-following responses to voice pitch in infants with normal hearing. Percept Mot Skills. 111, 765–784.

Johnson, K., Nicol, T., Zecker, S.G., Kraus, N., 2007. Auditory brainstem correlates of perceptual timing deficits. J Cogn Neurosci. 19, 376–385.

Johnson, K.L., Nicol, T., Zecker, S.G., Kraus, N., 2008. Developmental plasticity in the human auditory brainstem. J Neurosci. 28, 4000–4007.

King, C., Warrier, C.M., Hayes, E., Kraus, N., 2002. Deficits in auditory brainstem encoding of speech sounds in children with learning problems. Neurosci Lett. 319, 111–115.

Klatt, D., 1976. Software for cascade/parallel formant synthesizer. J Acoust Soc Am. 67, 971–975.

Kraus, N., McGee, T., Carrell, T.D., Zecker, S.G., Nicol, T.G., Koch, D.B., 1996. Auditory Neurophysiologic Responses and Discrimination Deficits in Children with Learning Problems. Science 273, 971–973.

Kraus, N., Slater, J., Thompson. E.C., Hornickel, J., Strait, D.L., Nicol, T., White-Schwoch, T., 2014. Auditory learning through active engagement with sound: Biological impact of community music lessons in at-risk children. Front Neurosci. 8, 1-12.

Kraus, N., White-Schwoch, T., 2015. Unraveling the Biology of Auditory Learning: A Cognitive-Sensorimotor-Reward Framework. Trends Cogn Sci. 19, 642-654.

Kraus, N., White-Schwoch, T., 2016. Newborn Hearing Screening 2.0. Hear J. 69, 44-45.

Krishnan, A., Swaminathan, J., Gandour, J., 2009. Experience-dependent enhancement of linguistic pitch representation in the brainstem is not specific to a speech context. J Cogn Neurosci. 21, 1092–1105.

Krishnan, A., Xu, Y., Gandour, J.T., Cariani, P.A., 2005.Encoding of pitchin the human brainstem is sensitive to language experience. Brain Res Cogn Brain Res. 25, 161–168.

Krizman, J., Slatera, J., Skoe, E., Marian, V., Kraus, N., 2015. Neural processing of speech in children is influenced by extent of bilingual experience. Neurosci Lett. 585, 48-53.

Lee, K.M., Skoe, E., Kraus, N., Ashley, R., 2009. Selective subcortical enhancement of musical intervals in musicians. J Neurosci. 29, 5832-5840.

Liu, F., Maggu, A.R., Lau, J., Wong, P., 2015. Brainstem encoding of speech and musical stimuli in congenital amusia: evidence from Cantonese speakers. Front Hum Neurosci. 8, 1-19.

Lyon, G.R., 1998. Why reading is not a natural process. Educ Leadersh. 55, 14-18.

Moeller, M.P., White, K.R., Shisler, L., 2006. Primary Care Physicians' Knowledge, Attitudes, and Practices Related to Newborn Hearing Screening. Pediatrics 118, 1357-1370.

Moller, A., 1999. Neural mechanisms of BAEP. Electroencephalogr Clin Neurophysiol Suppl. 49, 27-55.

Musacchia, G., Sams, M., Skoe, E., Kraus, N., 2007. Musicians Have Enhanced Subcortical Auditory and Audiovisual Processing of Speech and Music. Proc Natl Acad Sci U S A. 104, 15894–15898.

National Institutes of Health, Department of Health and Human Services, National Institute of Child Health and Human Development, 2000. National Reading Panel. Teaching Children to Read: An Evidence-Based Assessment of the Scientific Research Literature on Reading and Its Implications for Reading Instruction. Washington, DC: US Government Printing Office.

Neef, N., Schaadt, G., Friederici, A., 2017. Auditory brainstem responses to stop consonants predict literacy. Clin Neurophysiol. 128, 484-494.

Otto-Meyer, S., Krizman, J., White-Schowoch, T., Kraus, N., 2018. Children with autism spectrum disorder have unstable neural responses to sound. Exp Brain Res. 5164-4.

Parbery-Clark, A., Anderson, S., Hittner, E., Kraus, N., 2012. Musical experience offsets age-related delays in neural timing. Neurobiol Aging. 33, 1483.

Parbery-Clark, A., Skoe, E., Kraus, N., 2009. Musical experience limits the degradative effects of background noise on the neural processing of sound. J Neurosci. 29, 14100-14107.

Russo, N., Nicol, T., Musacchia, G., Kraus, N., 2004. Brainstem responses to speech syllables. Clin Neurophysiol. 115, 2021 – 2030.

Russo, N., Nicol, T.G., Trommer, B., Zecker, S., Kraus, N., 2009. Brainstem transcription of speech is disrupted in children with autism spectrum disorders. Dev Sci. 12, 557–567.

Russo, N., Nicol, T.G., Zecker, S.G., Haves, E.A., Kraus, N., 2005. Auditory training improves neural timing in the human brainstem. Behav Brain Res. 156, 95-103.

Schatschneider, C., Torgesen, J.K., 2004. Using our current understanding of dyslexia to support early identification and intervention. J Child Neurol. 19, 759 –765.

Skoe, E., Kraus, N., 2010. Auditory brainstem response to complex sounds: a tutorial. Ear Hear. 31, 302 – 324.

Skoe, E., Kraus, N., 2013. Musical training heightens auditory brainstem function during sensitive periods in development. Front Psychol. 4, 1-15.

Skoe, E., Krizman, J., Anderson, S., Kraus, N., 2013. Stability and plasticity of auditory brainstem function across the lifespan. Cereb Cortex. 25, 1415–1426.

Skoe, E., Nicol, T., Kraus, N., 2011. Cross-phaseogram: Objective neural index of speech sound differentiation. J Neurosci Methods. 196, 308-317.

Song, J.H., Skoe, E., Banai, K., Kraus, N., 2012. Training to Improve Hearing Speech in Noise: Biological Mechanisms. Cereb Cortex. 22, 1180-1190.

Song, J.H., Skoe, E., Wong, P.C., Kraus, N., 2008. Plasticity in the adult human auditory brainstem following short-term linguistic training. J Cogn Neurosci. 20, 1892–1902.

Strait, D. L., Kraus, N., Skoe, E., Ashley, R., 2009. Musical experience and neural efficiency: effects of training on subcortical processing of vocal expressions of emotion. Eur J Neurosci. 29, 661–668.

Strait, D.L., Kraus, N., 2014. Biological impact of auditory expertise across the life span: musicians as a model of auditory learning. Hear Res. 308, 109–121.

Strait, D.L., O'Connell, S., Parbery-Clark, A., Kraus, N., 2014. Musicians' enhanced neural differentiation of speech sounds arises early in life: developmental evidence from ages 3 to 30. Cereb Cortex. 24, 2512-2521.

Strait, D.L., Parbery-Clark, A., Hittner, E., Kraus, N., 2012. Musical training during early childhood enhances the neural encoding of speech in noise. Brain Lang. 123, 191–201.

Strait, D.L., Parbery-Clark, A., O'Connell, S., Kraus, N., 2013. Biological impact of preschool music classes on processing speech in noise. Dev Cogn Neurosci. 6, 51-60.

Stuart, A., Yang, E., Botea, M., 1996. Neonatal auditory brainstem responses recorded from four electrode montages. J Commun Disord 29, 125 – 139.

Stuart, A., Yang, E., Green, W., 1994. Neonatal Auditory Brainstem Response Thresholds to Air- and Bone-Conducted clicks: 0 to 96 Hours Postpartum. J Am Acad Audiol. 5, 163 – 172.

Vellutino, F.R., Fletcher, J.M., Snowling, M.J., Scanlon, D.M., 2004. Specific reading disability (dyslexia): what have we learned in the past four decades? J Child Psychol Psychiatry. 45, 2– 40.

Weiss, M.W., Bidelman, G.M., 2015. Listening to the brainstem: musicianship enhances intelligibility of subcortical representations for speech. J Neurosci. 35, 1687-1691.

Welch, P.D., 1967. The use of Fast fourier Transform for the estimation of power spectra: a method based on time averaging over short, modified periodograms. IEEE Trans. Audio Electroacoust. 15, 70-73.

White-Schwoch, T., Davies, E.C., Thompson, E.C., Woodruff Carr, K., Nicol, T., Bradlow, A.R. Kraus, N., 2015b. Auditory-neurophysiological responses to speech during early childhood: Effects of background noise. Hear Res. 328, 34-47.

White-Schwoch, T., Kraus, N., 2013. Physiologic discrimination of stop consonants relates to phonological skills in pre-readers: a biomarker for subsequent reading ability? Front Hum Neurosci. 7, 1-9.

White-Schwoch, T., Woodruff Carr, K., Thompson, E.C., Anderson, S., Nicol, T., Bradlow, A.R., Zecker, S.G., Kraus, N., 2015a. Auditory processing in noise: A preschool biomarker for literacy. PLoS Biol. 13, e1002196.

Wible, B., Nicol, T., Kraus, N., 2004. Atypical brainstem representation of onset and formant structure of speech sounds in children with language-based learning problems. Biol Psychol. 67, 299–317.

Willcutt, E.G., Pennington, B.F., 2000. Psychiatric comorbidity in children and adolescents with reading disability. J Child Psychol Psychiatry. 41, 1039 –1048.

Wong, P.C., Skoe, E., Russo, N.M., Dees, T., Kraus, N., 2007. Musical experience shapes human brainstem encoding of linguistic pitch patterns. Nat Neurosci. 10, 420–422.

Xu, Y., Krishnan, A., Gandour, J.T., 2006. Specificity of experience-dependent pitch representation in the brainstem. NeuroReport 17, 1601–1605.

Zuk, J., Ozernov-Palchik, O., Kim, H., Lakshminarayanan, K., Gabrieli, J.D., Tallal, P., Gaab, N., 2013. Enhanced syllable discrimination thresholds in musicians. PLoS One 8: e80546.

Figure legends

Fig. 1 Distribution of Wave V parameters. Scatter plots depicting all tested newborns (light color filled circles) within box plots for (**A**) latency, and (**B**) amplitude. In the scatter plots, dark filled circle and dark filled triangle indicate data corresponding to the individual newborns selected for illustration in successive figures, with high and low SNR respectively. In each box plot, thick black line and black filled diamond indicate the median and mean, respectively.

Fig. 2 Recording setup. Four disposable snap Ag/AgCl electrodes were placed in a vertical montage: the active electrode was located at Fpz (white electrode), references at each ear (red and blue, right and left respectively), and ground electrode at the forehead (black electrode). Only the ipsilateral reference (right electrode) was used in the analysis. The reproduction of the participant's picture is with the written consent of her mother.

Fig. 3 Spectral and temporal neural representation of the syllables /da/ and /ga/ in the newborn's auditory brain. (**A**) Stimulus waveforms (/da/, /ga/) plotted in black, and grand averaged time-domain FFR waveforms and amplitude FFR spectra extracted from the consonant transition and the vowel stimulus regions to /da/ and /ga/ ploted in red and blue, respectively. (**B**) Time-domain FFR waveforms and amplitude FFR spectra from an individual newborn with high SNR (NF 029) and (**C**) from another newborn with low SNR (NF 038).

Fig. 4 Spectrograms from FFRs elicited to syllables /da/ and /ga/. (**A**) Spectrograms extracted from the grand-averaged FFRs. (**B**) and (**C**) spectrograms corresponding to the same individual newborns with high and low SNR depicted in Fig. 3. The color scale from black to white indicates the spectral amplitude in µV, with light colours depicting highest amplitude values (top right).

Fig. 5 Distribution of each of the objective indices retrieved in the time domain. Scatter plots depicting all tested newborns (light color filled circles) within box plots for (**A**) stimulus-to-response cross-correlation; (**B**) neural lag; (**C**) signal-to-noise-ratio (SNR) of the consonant transition and (**D**) of the vowel; (**E**) pitch error; and (**F**) pitch strength measurements to /da/ and /ga/ syllables (depicted in red and blue, respectively). In the scatter plots, dark filled circles and dark filled triangles indicate data corresponding to the individual newborns with high and low SNR, respectively, as depicted in Fig. 3. In each box plot, the thick black line and the black filled diamond indicate the median and mean, respectively.

Fig. 6 Pitch tracking and points below noise floor obtained from (**A**) the grand average; (**B**) and (**C**) obtained from the individual newborns depicted in Fig. 3 with high and low SNR, respectively. F_0s were extracted separately from the FFRs to /da/ (*left column*) and /ga/ (*right column*), as well as from the stimuli (*black solid lines*) in overlapping steps of 40 ms using an autocorrelation based approach. Light grey areas indicate the bins in which the F_0 spectral amplitude was lower than the F_0 spectral amplitude in the pre-stimulus region (points below noise floor).

Fig. 7 Pitch strength obtained from (**A**) the grand averaged FFRs; (**B**) and (**C**) from the individual newborns depicted in Fig. 3 with high and low SNR, respectively. The figures depict the autocorrelation values (*r*; from -1 in *black* to 1 in *white*) computed separately from the FFRs to /da/ (*left column*) and /ga/ (*right column*) as a function of time and lag. The maximum autocorrelation value per unit of time is marked in black. Notice that the emerging black line is equivalent to the pitch tracking profiles observed in the Fig. 6.

Fig. 8 Distribution of each of the objective indices retrieved in the frequency domain. Scatter plots depicting all tested newborns within box plots for: (**A**) F_0 and averaged $HH_{2\text{-}13}$ spectral

amplitudes during the consonant transition, and (**B**) the vowel; and (**C**) points below noise floor measurements, all computed for each syllable separately. The layout is equal to that used in Fig. 5.

706 **Table 1**

707 Time-domain parameters. Descriptive statistics of stimulus-to-response cross-correlation,
708 neural lag, RMS amplitude calculated from the pre-stimulus region, SNR computed from the
709 consonant transition, vowel and entire response, pitch error and pitch strength for each syllable
710 presented.

711

Measure	**/da/** **M (SD)**	**/ga/** **M (SD)**
Cross-correlation (Pearson's r)	0.15 (0.04)	0.15 (0.03)
Neural lag (ms)	5.87 (1.11)	5.90 (1.16)
Pre-stimulus RMS (μV)	0.03 (0.01)	0.03 (0.01)
SNR		
Consonant transition (10 – 57 ms)	1.58 (0.75) †	1.82 (0.66)
Vowel (57 – 170 ms)	1.48 (0.41)	1.44 (0.48) †
Entire response (0 – 170 ms)	1.53 (0.44)	1.46 (0.69) †
Pitch error (Hz)	4.89 (2.01)	5.29 (1.78)
Pitch strength (r)	0.61 (0.18)	0.63 (0.17)

712 † Median (IQR, interquartile range)

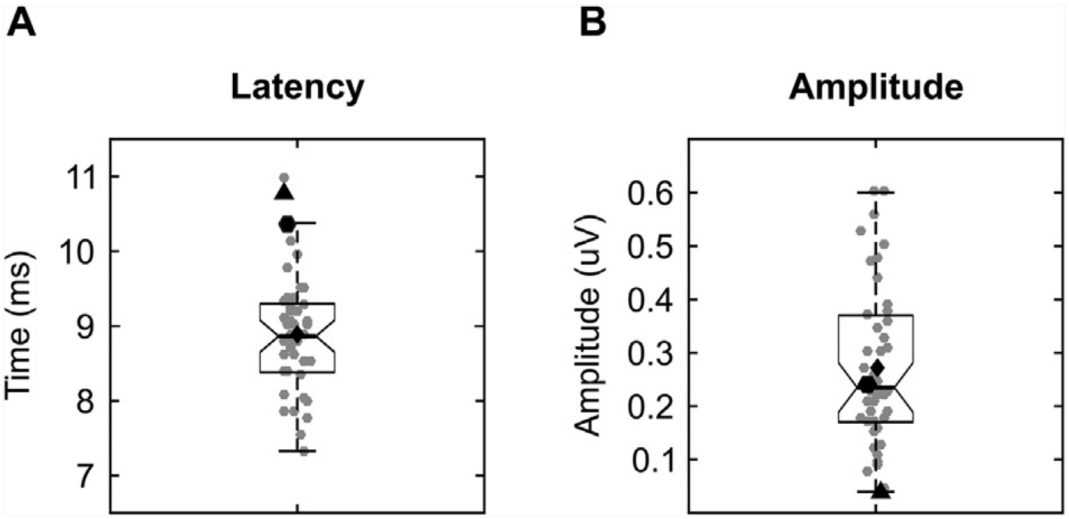
Table 2

Frequency-domain parameters. Descriptive statistics for spectral amplitude across response syllables, stimulus regions and harmonics and for points below noise floor for each syllable presented.

Measure	**/da/** **M (SD)**	**/ga/** **M (SD)**
Spectral Amplitude (nV)		
Consonant transition (10 – 57 ms)		
F_0	40.65 (27.82) †	44.01 (21.45)
$HH_{2\text{-}13}$	3.01 (0.86)	2.52 (0.99) †
Vowel (57 – 170 ms)		
F_0	35.43 (13.55) †	34.47 (14.42)
$HH_{2\text{-}13}$	2.88 (1.16) †	2.73 (0.66)
Entire response (0 – 170 ms)		
F_0	38.29 (13.98)	38.11 (14.84)
$HH_{2\text{-}13}$	3.19 (0.89)	2.83 (0.65)
Points below noise floor	2.5 (24.75) †	0 (13.25) †

† Median (IQR, interquartile range)

721 **Figure 1**
722



723
724

725 **Figure 2**
726

727

Figure 3

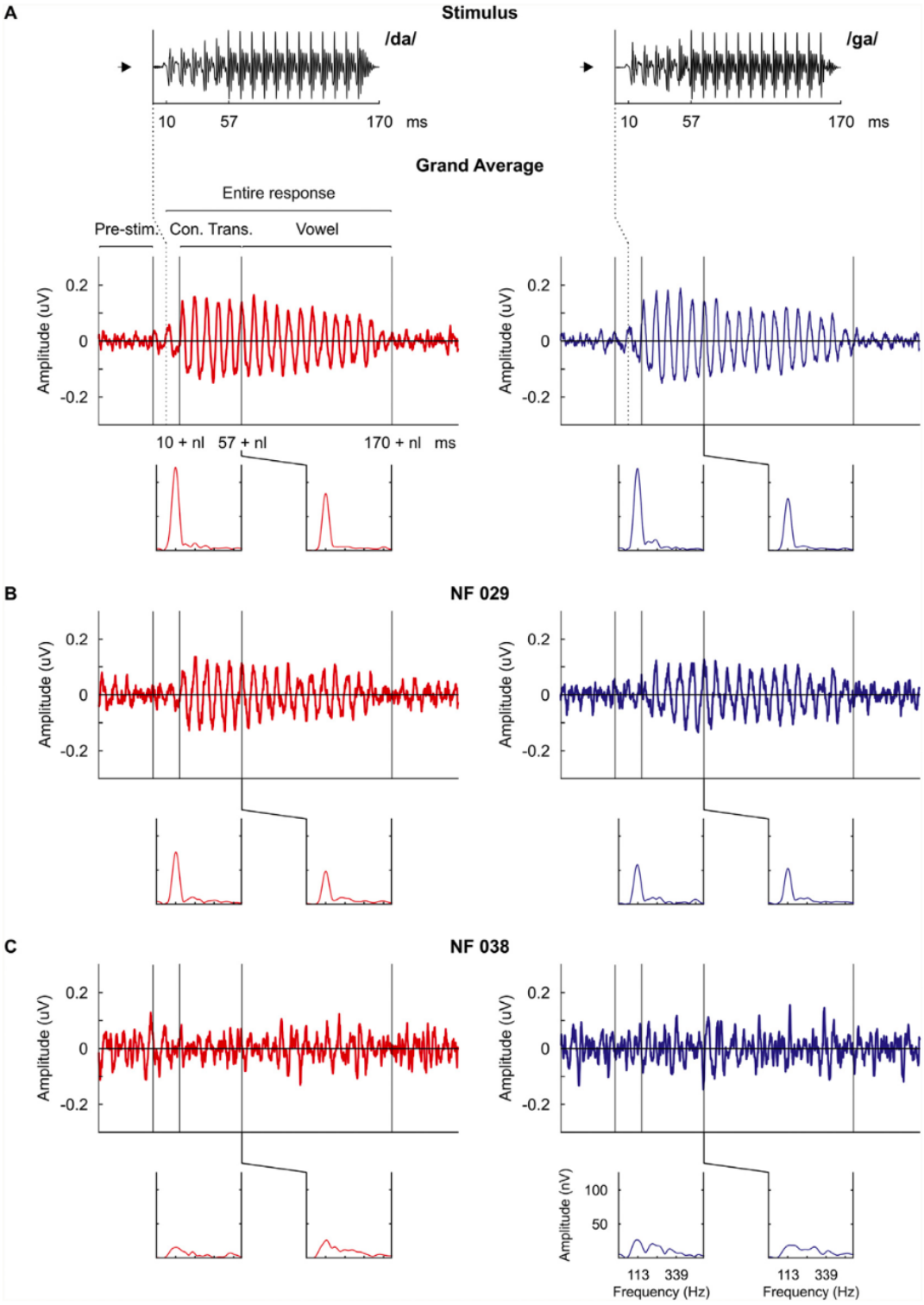

A
Stimulus
/da/
/ga/
10
57
170 ms
Grand Average
Entire response
Pre-stim.
Con. Trans.
Vowel
Amplitude (uV)
0.2
0
-0.2
10 + nl
57 + nl
170 + nl ms
B
NF 029
C
NF 038
Amplitude (nV)
100
50
113
339
Frequency (Hz)

Figure 4

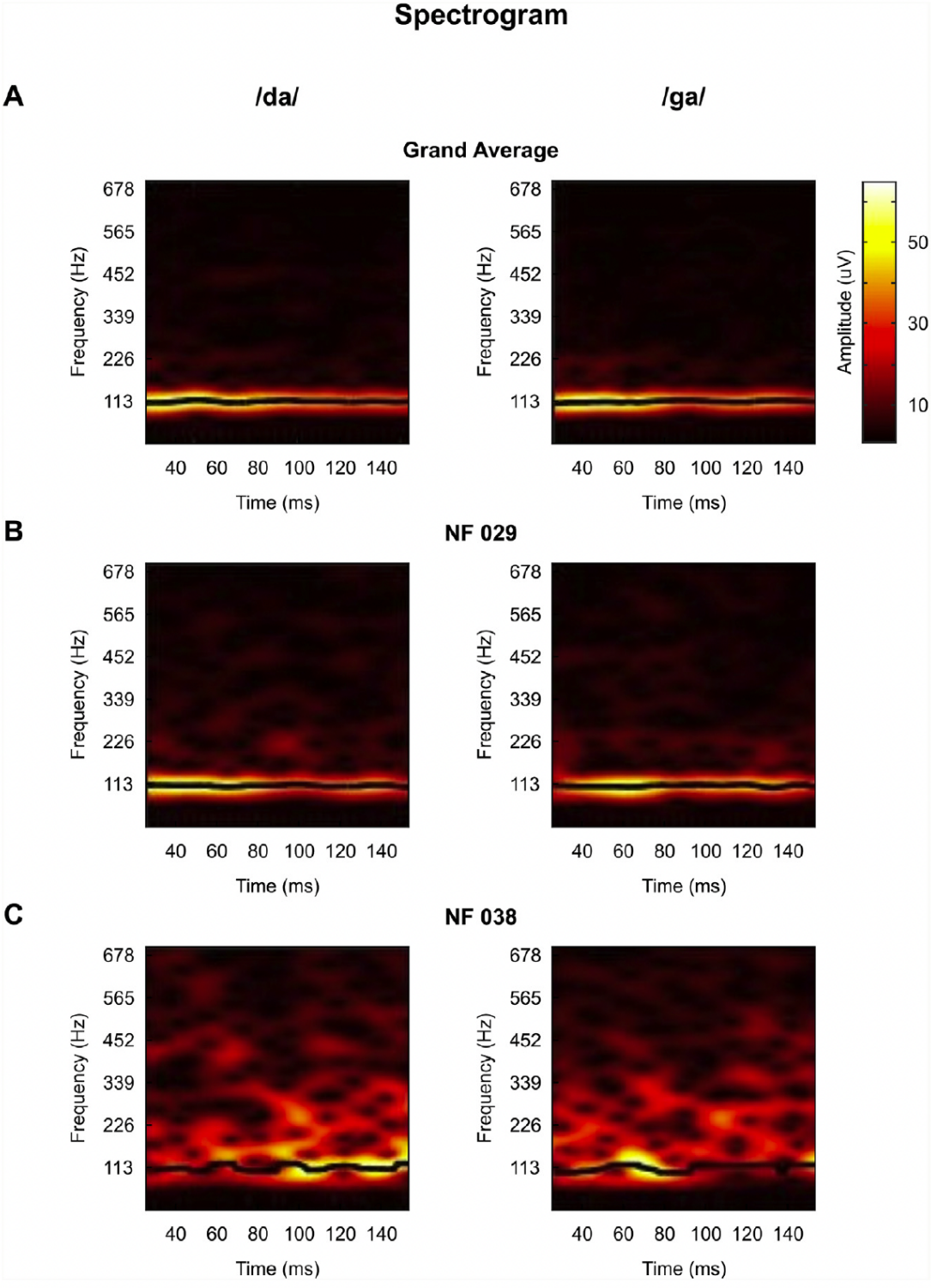

Spectrogram
A
/da/
/ga/
Grand Average
B
NF 029
C
NF 038
Frequency (Hz)
678
565
452
339
226
113
40 60 80 100 120 140
Time (ms)
Amplitude (uV)
50
30
10

Figure 5

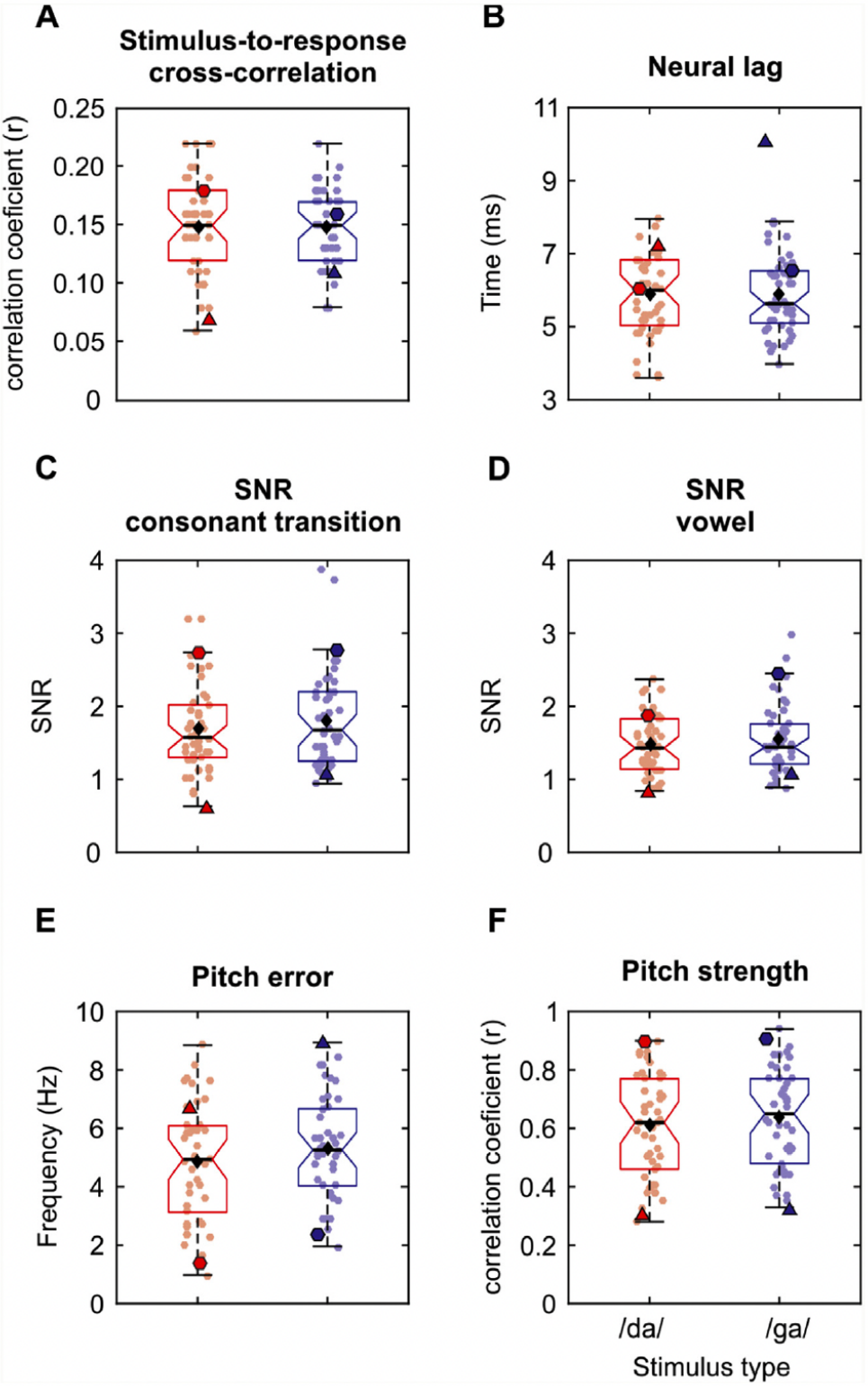
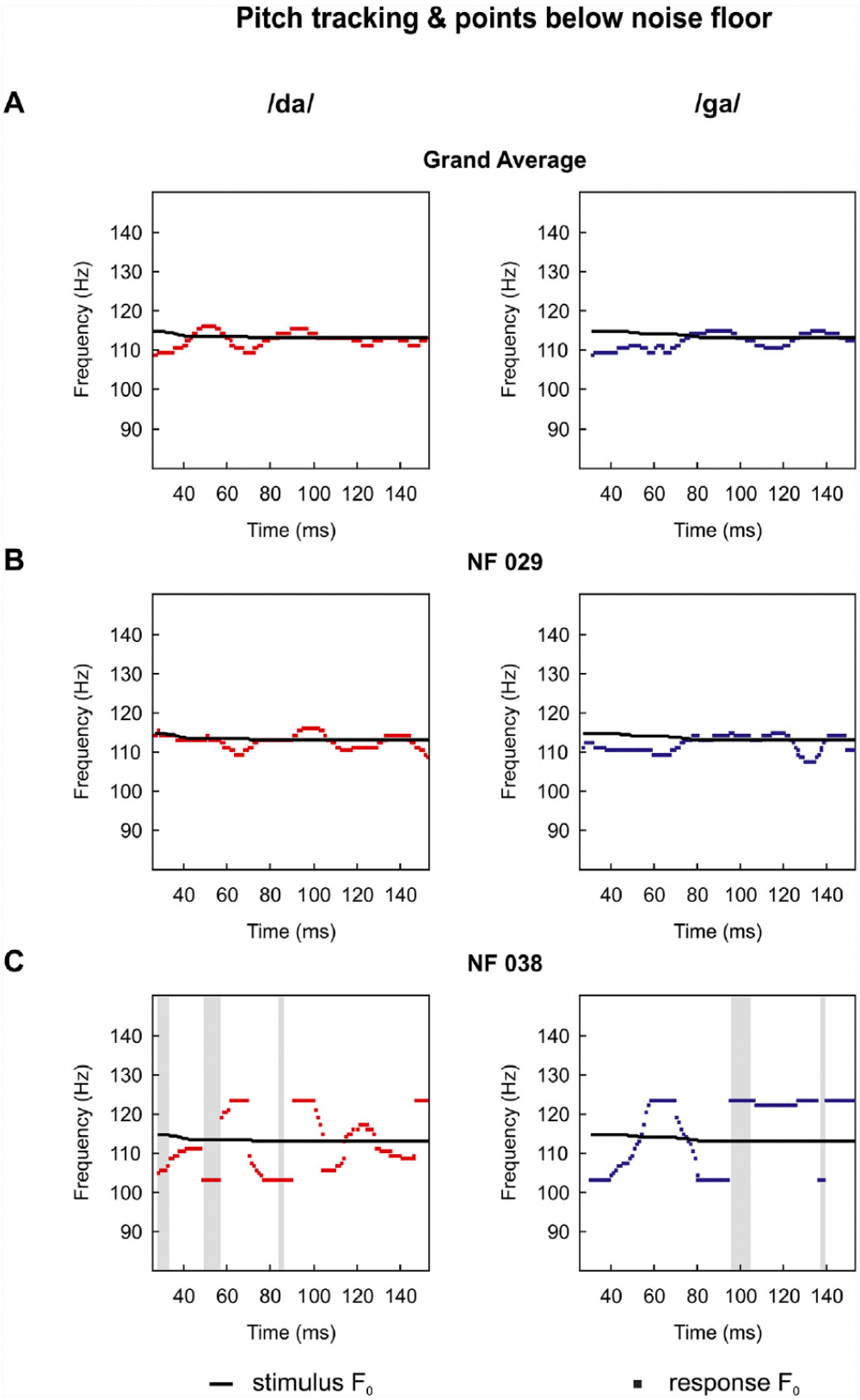
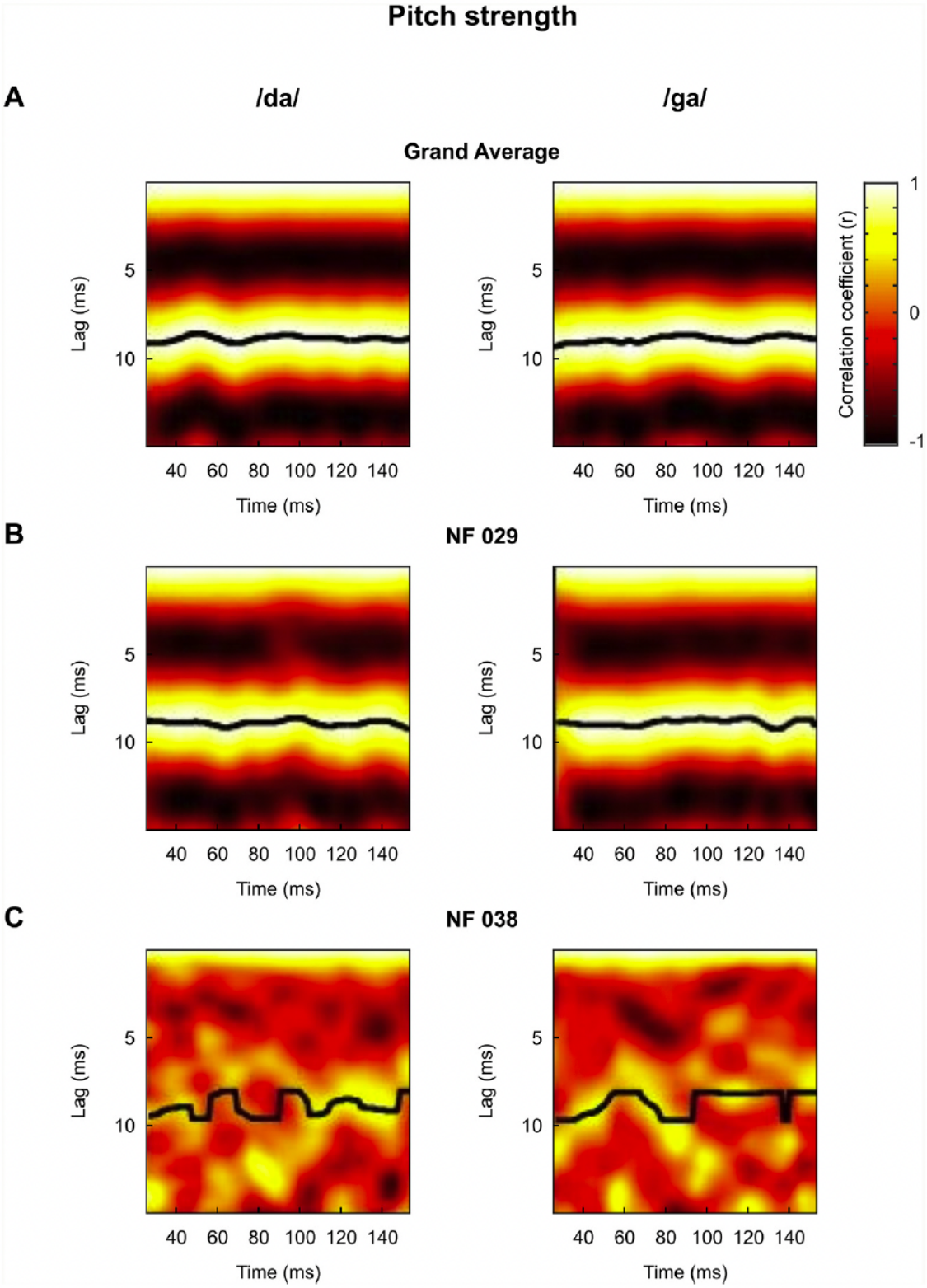

A
Stimulus-to-response cross-correlation
correlation coeficient (r)
0.25
0.20
0.15
0.10
0.05
0
B
Neural lag
Time (ms)
11
9
7
5
3
C
SNR consonant transition
SNR
4
3
2
1
0
D
SNR vowel
SNR
4
3
2
1
0
E
Pitch error
Frequency (Hz)
10
8
6
4
2
0
F
Pitch strength
correlation coeficient (r)
1
0.8
0.6
0.4
0.2
0
/da/
/ga/
Stimulus type

Figure 6


Pitch tracking & points below noise floor
A
/da/
/ga/
Grand Average
Frequency (Hz)
Time (ms)
B
NF 029
C
NF 038
stimulus F0
response F0

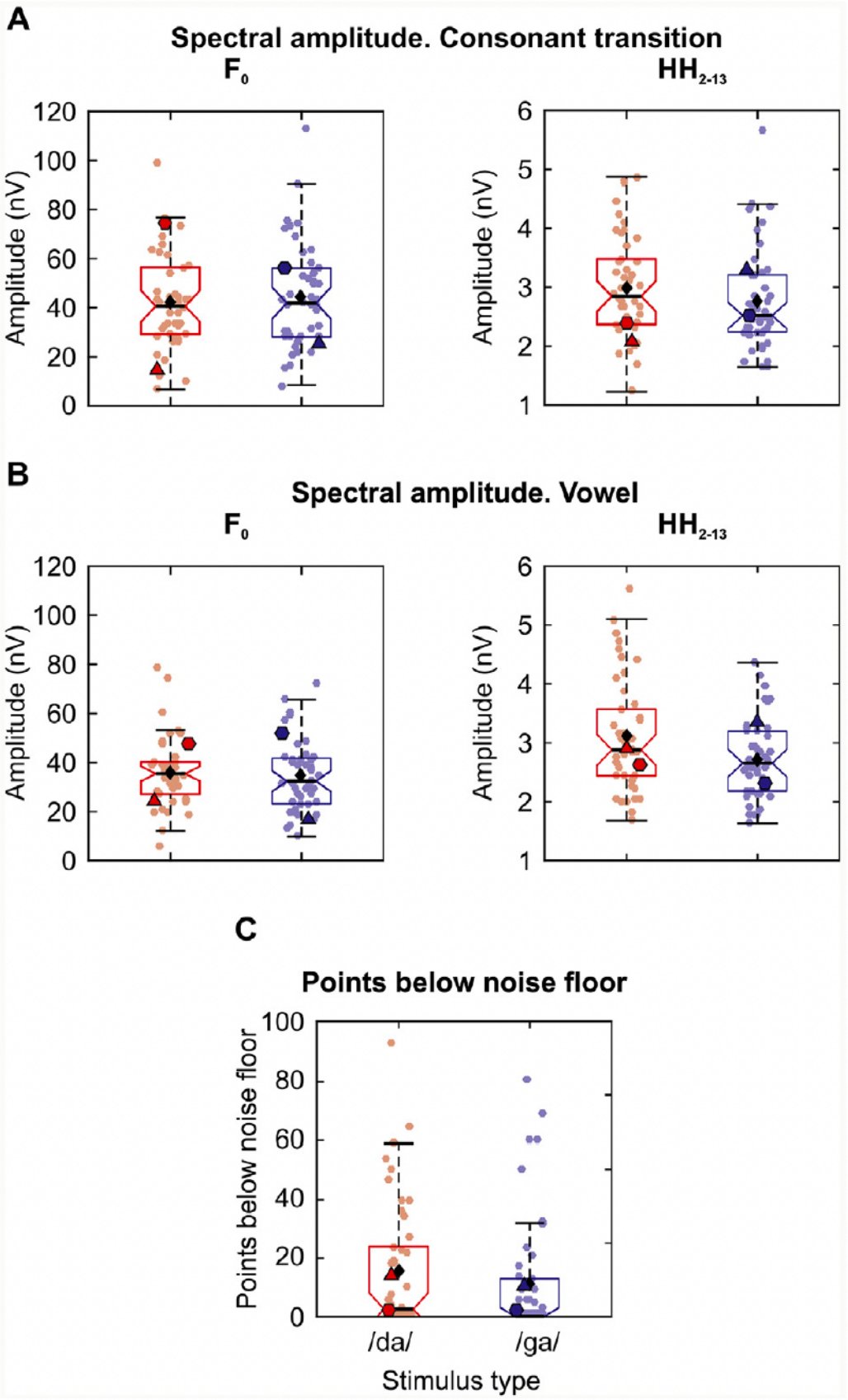
742 **Figure 7**

743



744

745 **Figure 8**
746



747