Clinical Neurophysiology 129 (2018) 2623–2634

Contents lists available at ScienceDirect

Clinical Neurophysiology

journal homepage: www.elsevier.com/locate/clinph

# Effects of noise and age on the infant brainstem response to speech

Gabriella Musacchia [a,c,*], Silvia Ortiz-Mantilla [b], Cynthia P. Roesler [b], Sree Rajendran [b], Julie Morgan-Byrne [b], April A. Benasich [b]

[a] Department of Audiology, University of the Pacific, 155 Fifth Street, San Francisco, CA 94103, United States
[b] Center for Molecular and Behavioral Neuroscience, Rutgers University, Newark, NJ, United States
[c] Department of Otolaryngology Head and Neck Surgery, Stanford Medical School, Stanford, CA, United States

## ARTICLE INFO

*Article history:*
Accepted 24 August 2018
Available online 30 August 2018

*Keywords:*
Infant
Frequency following response (FFR)
Auditory brainstem response (ABR)
Evoked potentials
Speech in noise
cABR

## HIGHLIGHTS

- Older infants (18–24 mo) have more robust speech in noise brainstem responses than younger infants (7–12 16 mo).
- Responses to clicks and speech in quiet did not differ across age.
- Speech-in-noise responses may enable early identification of functional auditory processing problems.

## ABSTRACT

*Objective:* Background noise makes hearing speech difficult for people of all ages. This difficulty can be exacerbated by co-occurring developmental deficits that often emerge in childhood. Sentence-type speech-in-noise (SIN) tests are available clinically but cannot be administered to very young individuals. Our objective was to examine the use of an electrophysiological test of SIN, suitable for infants, to track developmental trajectories.
*Methods:* Speech-evoked brainstem potentials were recorded from 30 typically-developing infants in quiet and +10 dB SNR background noise. Infants were divided into two age groups (7–12 and 18–24 months) and examined across development. Spectral power of the frequency following response (FFR) was computed using a fast Fourier Transform. Cross-correlations between quiet and noise responses were computed to measure encoding resistance to noise.
*Results:* Older infants had more robust FFR encoding in noise and had higher quiet-noise correlations than their younger counterparts. No group differences were observed in the quiet condition.
*Conclusions:* By two years of age, infants show less vulnerability to the disruptive effects of background noise, compared to infants under 12 months.
*Significance:* Speech-in-noise electrophysiology can be easily recorded across infancy and provides unique insights into developmental differences that tests conducted in quiet may miss.

## 1. Introduction

Speech perception depends on healthy sound transduction and faithful encoding of speech sound acoustics. Background noise can distort frequency analysis in the inner ear and disrupt auditory processing more centrally, making speech sounds difficult to decode. Problems in the middle ear, inner ear or central nervous system can exacerbate difficulties in hearing speech in noise, including early conductive hearing impairment (Keogh et al., 2010), sensorineural hearing loss at low or high frequencies (Laukli and Mair, 1985; Horwitz et al., 2002), diminished linguistic content during development (Cooper and Aslin, 1989; Lieu, 2004; Stelmachowicz et al., 2004; Eisenberg, 2007; Moeller and Tomblin, 2015), and attention (Soderlund and Jobs, 2016) or memory problems (McCreery et al., 2017; Millman and Mattys, 2017). Because of this, hearing in noise can be especially difficult for people with hearing loss (Brons et al., 2014), those with language-learning deficits and delays (Bradlow et al., 2003; Sperling et al., 2005; Ziegler et al., 2011; Vance and Martindale, 2012), older adults (Helfer and Freyman, 2008; Moore et al., 2014), infants

---

* Corresponding author at: Department of Audiology, University of the Pacific, 155 Fifth Street, San Francisco, CA 94103, United States.
*E-mail address:* gmusacchia@pacific.edu (G. Musacchia).

https://doi.org/10.1016/j.clinph.2018.08.005
1388-2457/

(Leibold et al., 2016) and young children (Jamieson et al., 2004; Leibold and Buss, 2013). Even in children and young adults with audiograms within normal clinical limits, at least 5% find speech sounds less intelligible in noisy environments (Hind et al., 2011). Despite this prevalent problem, objective measures of "speech in noise" processing are only beginning to be explored. In this study, we show, for the first time, how noise affects the infant brainstem response to speech over development. Our neurophysiological results suggest that noise impacts the developing auditory system differently as humans age.

### 1.1. Speech in noise perception

Speech-in-noise tests are typically administered in an Audiology clinic to quantify how much an individual's speech perception is disrupted by noise. Although these tests use sentence stimuli to assess global hearing in noise, research has shown that not all acoustic components of speech are affected equally by noise. Rapid, low-amplitude acoustic shifts, formed by the stoppage of airflow during stop consonants, are easily susceptible to disruption by the random acoustic occurrences in background noise (Brandt and Rosen, 1980). Longer-lasting, periodic portions of speech, such as vowel sounds like [a] or [e], are less likely to be disrupted by the random effects of noise because they are sustained over longer periods of time, which allows more encoding opportunities and spectral grouping over time (Miller and Jusczyk, 1989).

One of the most salient cues in vowel sounds is "voice pitch", which is the perceptual cue related to the spectrum of the sound, especially the fundamental frequency (F0). The F0 of a speech sound corresponds to the glottal-pulse rate, changes with pitch contour, and contributes to the perception of prosody, voicing (Faulkner and Rosen, 1999), lexical segmentation (Spitzer et al., 2007) and speaker identity (Smith and Patterson, 2005; Baumann and Belin, 2010). Frequencies above the F0, called harmonics, define the type of vowel (e.g. "ah" vs "oo") are created by the shape of the oral cavity during voicing. Spectral masking, or smearing, impacts children's speech perception more than adults; doubling the errors in noise (Nittrouer et al., 2015) and can produce two-fold threshold elevations compared to adults during signal detection tasks (Allen and Wightman, 1995; Oh et al., 2001; Leibold and Neff, 2011). In impaired listeners, such as cochlear implant recipients, effects of masking the F0 can be even greater (Qin and Oxenham, 2003).

While strong advances in speech-in-noise testing have been made over the past decade, current clinical tests of perception do not distinguish between F0 and consonant difficulties and are constrained to measures requiring a patient's behavioral response. These limitations can decrease tester objectivity, reduce test specificity and preclude testing of pre–verbal infants or non-verbal patients who may benefit from identification of speech-in-noise deficits. This gap in our knowledge prevents Audiologists and associated medical professionals from detecting signs of possible language problems in nonverbal patients, considering specific remediation strategies, and counseling patients or parents for follow-up testing and/or early intervention for pre-verbal infants. Thus, an objective, sensitive and reliable method to test the specific effects of noise that can be used during development is needed to fill in these gaps.

### 1.2. Neural correlates of speech-in-noise perception

Research over the past 15 years has produced abundant evidence establishing the complex Auditory Brainstem Response (cABR), and in particular, the Frequency Following Response (FFR), as an objective and reliable measure of speech processing across the lifespan [for review, see (Skoe et al., 2015; Skoe and Kraus, 2010; Kraus et al., 2017)]. The cABR is a single-channel electrophysiological test that records the brain's response to the syllable /da/, usually in under 20 min. The physical setup of the recording is identical to the well-established Auditory Brainstem Response (ABR) test, in which Audiologists measure a patient's brain response to a click stimulus in order to estimate audiometry or neurophysiological integrity (Hall, 1992; Hall and Rupp, 1997). Although the click-evoked response is considered the gold standard measure for demonstrating clinical abnormalities in many disorders, it reflects the broadband response of the auditory system between 2 and 4 kHz and is dominated by firing of neurons that respond best to sound onset. Therefore, the ABR response is only able to identify serious deficits that undermine the basic integrity of the auditory system in patients with suspected hearing loss and certain types of vestibular disorders.

The cABR stimulus on the other hand, was designed to provide fine-grained measures of phonetic information processing that clinical populations have particular difficulty perceiving [e.g. (Russo et al., 2004; Kraus and Nicol, 2005)]. The cABR response mechanisms can be broadly divided into two parts: (1) transient response mechanisms that encode consonant onset ("d'), similar to the click-evoked response described above, and ultra-rapid frequency shifts, similar to a chirp or fast frequency sweep, and (2) the sustained vowel portion ("a"), which entrains to the periodicities present in the stimulus via phase-locked intervals occurring at periods of the F0 at ~100 Hz. The sustained response elicited by the cABR vowel is more generally called the FFR, which can be elicited by any periodic stimulus up to ~1000 Hz (Sohmer et al., 1977; Hoormann et al., 1992; Krishnan et al., 2005) and is strongest when speech is intelligible (Galbraith et al., 1995, 1997, 2004). In the case of the cABR, the transient acoustics of the consonant elicit an onset response characterized by a positive (wave V) and negative (wave A) going peak occurring at ~7.5 ms. The response to the sustained F0 follows the onset response, and is characterized by waves D, E, and F, each separated by a period of ~10 ms. Primary generators of the cABR and more generally, the FFR, have been localized to the fluctuation of the endolymph at the apex of the cochlear hair cells and the phase-locked excitatory post-synaptic potentials of neurons in the inferior colliculus [for review, see (Bhagat, 2012)].

The cABR has been used to assess speech processing in preschoolers (Johnson et al., 2005), school-aged children (Russo et al., 2004, Hornickel et al., 2012), adults (Krishnan et al., 2005; Song et al., 2011a) and the elderly (Anderson et al., 2012). Because of its ability to capture complex auditory processing mechanisms, measures of the cABR have been able to show processing differences in language- and reading-impaired populations where other more gross measures of auditory processing, such as the click-evoked ABR, do not (King et al., 2002; Wible et al., 2004; Abrams et al., 2006). Data from these and other studies (Hornickel et al., 2009; Anderson et al., 2013) show that the cABR and FFR is more impacted by noise in impaired populations, compared to normal listeners. This makes the cABR and FFR clinically relevant for those seeking an objective measure of hearing in noise.

How does noise impact the brainstem response to speech? One of the first reports of the noise effect (Russo et al., 2004) showed maximal disruption in the onset response to the consonant in the syllable /da/ and less degradation of the periodic F0. In that study, noise delayed onset peak latencies by 0.5–0.9 ms and diminished peak amplitudes by 74–92% effectively distorting the timing and magnitude of the signal. It is important to note here, that the timing of the brainstem response is extremely precise and reliable in individuals, rendering timing differences of this order significant in both the laboratory and clinic. Noise also affected the sustained vowel periodicity but to a lesser degree; degrading the spectral representation of the F0 by about ~30%. The distortion of timing and reduction of the F0 shown in that study was hypothesized to contribute to difficulties decoding speech signals in noise. Subsequent reports from Song, Anderson and colleagues, replicated these findings and established a strong relationship between speech-in-noise perception and measures of the cABR in noise; suggesting that F0 encoding in noise is particularly degraded in people who have difficulty understanding speech in noise. The above data suggest that the cABR recorded in background noise may be a suitable and more sensitive tool for the investigation of functional speech processing.

### 1.3. The current study

Recently, the feasibility of recording the cABR in preverbal infants was demonstrated over 3–10 months-of-age (Anderson et al., 2015). In that cross-sectional study, the authors showed that representation of the speech F0 was robust, reliable and mature at 3 months, while higher frequency components of the FFR and peak latencies continued to mature across age. These results support animal models demonstrating earlier development for lower frequency responses in brainstem nuclei [e.g. (Rubel and Ryals, 1983)]. Taken together, the animal and infant data strongly suggest that maturation of spectral resolution for speech is specific to higher frequency features of sound in the first year of life. At this time it is unknown however, whether and to what extent the addition of noise impacts speech processing and thus, language acquisition, in infants.

In the current study, we begin to address the gap in early identification of speech-in-noise problems by evaluating the cABR in Quiet and +10 dB SNR Noise conditions in two infant populations: (1) younger infants ages 7–12 months (YI) and (2) older infants ages 18–24 months (OI). This time frame involves extensive development of the hearing system (Lenneberg, 1967), and particularly, at the central level (Eggermont, 1985). Many aspects of basic auditory processing are adult-like by the middle of the first year of life (Eggermont, 1985). However, processes requiring more complex processing computations, such as auditory attention, speech segmentation and localization, take longer to develop (Eilers et al., 1981; Eilers, 1985; Muir, 1985). The overall age group of 7–24 months-of-age was selected in order to span important developmental milestones of speech reception and production (Werker et al., 1981; Werker and Tees, 1983, 2005; Luinge et al., 2006). From 7 to 12 months of age, infants develop rudimentary complex auditory skills such as turning to the location of sounds, imitating speech and recognizing some words. In addition, in this highly plastic developmental period, young infants narrow their phonemic perceptual abilities, become attuned to their native language and build cortical representations of the familiar sounds in their surrounding linguistic environment (Werker and Hensch, 2015; Ortiz-Mantilla et al., 2016). Over the span of 18–24 months of age, as phonemic mapping is consolidated in the auditory cortex, speech and language skills greatly improve, ending with the accelerating acquisition of new words on a regular basis and older infants able to communicate or follow simple commands such as "roll the ball". Electrophysiologically, previous data from our laboratory has shown that auditory ERPs in the younger infant age range correlate strongly to ERP measures in the older infant age range (Benasich et al., 2006; Choudhury and Benasich, 2011). In addition, ERPs originating from the 6- and 9-month age range were predictive of language outcome at 3–4 years of age. Taking into account Anderson et al.'s results, we predicted that differences in the Quiet condition would be restricted to the high frequency (HF) range and that differences in Noise would be greatest in the HF frequency band. However, given the prominent role of voice pitch and the developmental trajectory of speech-in-noise, we also expected the OI group to show less degradation of the F0 in noise, compared to YI counterparts. To the authors' knowledge, the results of this study show for the first time that that the effects of noise change with development over the first year of life. In addition, our method underscores the feasibility of speech-in-noise response acquisition in very young populations.

## 2. Methods

### 2.1. Participants

This study was conducted in accordance with the Declaration of Helsinki. Informed consent, approved by the Institutional Review Board of Rutgers University, was obtained from all parents before study participation. Parents were compensated for their time, and infants received a toy after the visit. A total of thirty (30) infants participated in this study, with data from two infants being excluded due to SNR <1. All infants recruited were full-term, had normal birth-weight, uneventful pre and perinatal circumstances and did not present congenital, neurological or physical abnormalities or impairments. Exclusion criteria included family history of language impairments in the nuclear family (e.g. diagnosed autism, specific language impairment, dyslexia, attention deficit or hyperactivity disorder). Infants were separated into two groups by age with Older Infants ages 18–24 months (N = 14, mean age = 21.8 m, SD 3.03, 7 females) and Younger Infants ages 7–12 months (N = 14, mean age = 10.2 m, SD 2.07, 7 females).

### 2.2. Stimuli

Stimuli consisted of an acoustic click (100-μs square wave) and speech syllable, /da/. The click produced a broad frequency spectrum of approximately 2000–4000 Hz. The speech syllable (Fig. 1A) was created with Klatt-based software (Klatt, 1980) and consisted of a 40 ms computer-synthesized sound that mimics the properties of a human produced /da/. The consonant ("d") acoustics reflect stoppage of airflow and contained an initial five-millisecond burst of frequencies ranging from 2500 to 4500 Hz. This was followed by an acoustic transition to the sustained periodicity in which the F0 rose linearly from 103 to 125 Hz. Five formants in the high frequencies (HF) consisted of frequency shifts in three bands: 220–720 Hz (F1), 1700–1240 Hz (F2) and 2850–2500 Hz (F3). F4 and F5 stayed constant at 3600 Hz and 4500 Hz, respectively. Stimuli were presented monaurally via right insert earphone (ER-3, Etymotic Research, Elk Grove Village, IL) at 70 dB (calibrated with model 2250, Bruel and Kjaer, Germany 1996) using Intelligent Hearing Systems IHS5441 (Miami, FL USA).

### 2.3. Procedure and data recording

During testing, infants were seated on their caregiver's lap while an experimenter engaged the infant's attention with quiet play (puppets, bubbles, etc.). Age-appropriate movies or cartoons were also played at very low volume on a video monitor in front of the children. Responses were collected using a one-channel vertical montage with electrodes placed at Cz (active, non-inverting) to Ai (ipsilateral mastoid, reference) and Fpz (ground). Recording protocol and sweep number were as follows. Two blocks of 1024 clicks were presented before (PRE) and after (POST) the speech conditions. The click-evoked response was used as a baseline response at the beginning and end of the recording. The PRE click block was used to ensure that the infant did not have any gross neurophysiological deficits that could be identified according to the normative click-evoked ABR values. The POST click was used to estimate neuronal adaptation or shifts in neuronal function or experimental setup that may have occurred during the testing ses-

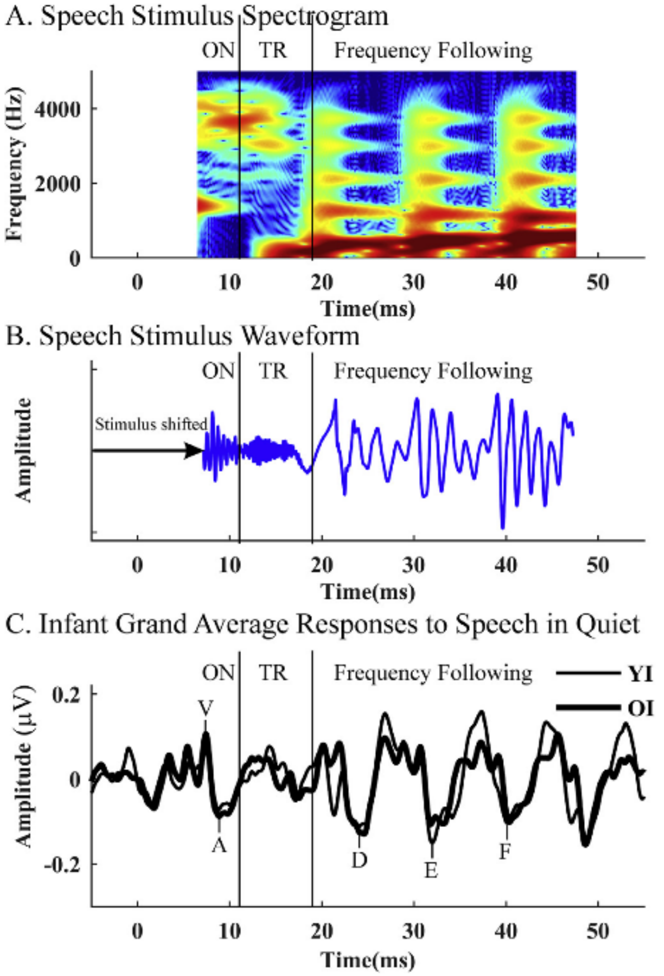

**Fig. 1.** Stimulus spectrogram, waveform and infant grand average brain responses in Quiet. (A) The 40 ms synthesized speech stimulus /da/ comprises a high frequency burst at the consonant, followed by a formant transition into sustained periodicity of the vowel. Acoustic features are separated into Onset (ON, 0–12.5 ms), Transition (TR, 12.5–20 ms) and the sustained Frequency Following Response (FFR, 20–45 ms). (B) Temporal features of each acoustic section can be observed in the stimulus waveform. The stimulus waveform is shifted forward in time by 7 ms to approximate neural transduction time from cochlea to brainstem. The period (T = ~10 ms) and fundamental frequency (F0 = ~100 Hz) of the vowel can be observed by measuring the time between the largest peaks. (C) Grand average brain responses to /da/ in Quiet for YI (7–12 months infants) and OI (18–24 months infants) mirror the acoustic features of the stimulus and are divided into ON, TR, and FFR sections. Waves V and A reflect encoding of the sound onset and typically occur before 8–10 ms post stimulus. Peaks D, E and F occur with a period of about 10 ms/peak and comprise FFR, beginning about 21 ms post stimulus.

sion. After the PRE click recording, one block of 2000 alternating polarity /da/ stimuli was presented Quiet and ipsilateral +10 dB SNR white Gaussian background noise. Online averaging settings included a recording time of a −20 ms baseline and 15 ms post-stimulus onset, a sampling rate of 20,000 Hz, an online bandpass filter of 100–1500 Hz and an artifact rejection criteria of ±35 μV.

### *2.4. Event-Related potential measurement and analysis*

Click-evoked ABR waveform peaks are conventionally labeled sequentially as waves I-V and occur within a 10-millisecond time period after sound onset. The most prominent feature of the click-evoked ABR, Wave V, is the primary means by which hearing threshold is evaluated in non-verbal individuals. In our study, Wave V was marked in both PRE and POST waveforms for each participant.

The speech-evoked cABR is a complex waveform that faithfully mimics the spectro-temporal fluctuations of the stimulus (Fig. 1). The response is conventionally divided into three acoustic components: (1) the short transient response speech onset, (2) the transition portion consisting of rapid frequency shifts, and (3) the sustained, periodic portion, comprised of responses to the F0 and HF that give rise to the frequency following response (FFR, Fig. 1C) (King et al., 2002; Russo et al., 2004; Wible et al., 2004; Johnson et al., 2005; Kraus and Nicol, 2005; Musacchia et al., 2007; Johnson et al., 2008; Hornickel et al., 2009; Skoe and Kraus, 2010). In general, the first two elements support perception of the stop consonant [d] and the third element contributes to perception of the vowel [a]. The F0 of the stimulus (~100 Hz) can be verified in the response waveform (Fig. 1C) by measuring the time between the largest peaks.

Latency and amplitude values of discrete peaks V, A, C, D, E and F were assessed to measure the timing of response (Fig. 1B and C). Four independent raters picked peaks V, A, D, E and F in the individual averages collected in the Quiet condition. In the noise condition, peaks were often degraded or obscured in the individual average. These peaks were designated as "CND" or Could Not Detect and the latency and amplitude were excluded from analysis. These exclusions constitute "detectability rates" and are described in the Results section. In cases of peak marking differences between raters that were greater than ±1 ms, a discussion among all investigators was held to determine the best peak to mark.

The FFR was defined by the sustained, periodic portion in which peaks D, E and F followed the periodicity of the stimulus F0. This corresponded to roughly 20–45 ms post stimulus onset (Fig. 1). Because the periodicity of discharges appears to "follow" the frequency of the sound, spectral analysis is often used to assess the robustness of the FFR. To assess spectral amplitude, a fast Fourier transform (FFT) analysis was computed over the FFR epoch of each individual average. Each individual FFT response was baseline corrected to the spectrum of the pre-stimulus period (−20 to 0 ms). The following five FFR analysis techniques were employed according to previously published parameters (Russo et al., 2004; Song et al., 2011a, 2011b). Maximum amplitude of the F0 frequency component (Max F0) was determined via custom MATLAB code over the range of 80–120 Hz. Root mean square amplitude (RMS) was calculated over this range (F0 RMS) to give a broader view of the F0 magnitude. RMS amplitude was also calculated over the spectral component corresponding to 200–720 Hz; the high frequencies present in the first formant (HF RMS). Stimulus-to-response (SR) correlations were calculated using the Signal Processing toolbox and custom MATLAB code (see xcorr.m, https://www.mathworks.com/help/signal/ref/xcorr.html) in the Quiet and Noise conditions. The SR correlation provides a measure of how faithfully the response mirrors the stimulus waveform and provides an index of neuronal synchrony, related to an individual's phase-locking capacity. The SR calculation results in two values: (1), the "lag" or duration of shift, and (2) the "r-value" which is a measure of the goodness of fit. An r-value of one (1) relates to complete correlation and a value of zero (0) relates to no correlation between the two signals. To obtain the r-value, each individual FFR response was cross-correlated to the 20–45 ms portion of the stimulus syllable. Only maximal correlations with a greater than five (5) ms lag were accepted to account for the conduction time of neural impulses. Both the lag and r-values were used as variables of interest in our statistical analyses.

### *2.5. Statistical analysis*

Pearson's correlation tests were conducted to determine relationships among brainstem measures in Quiet as reported in (Russo et al., 2004). A series of ANOVA tests assessed effects of age and noise on the two groups for latency and FFR measures. Post-hoc t-tests were conducted where appropriate and Least Significant Difference (LSD) correction was used for multiple comparisons.

## 3. Results

### 3.1. Stable click-evoked ABR within normative ranges

Click-evoked ABRs were recorded at the beginning (PRE) and end (POST) of each testing session. No significant difference was observed between PRE and POST latencies for either OI or YI infants ($p > 0.05$), suggesting temporal acuity was preserved throughout the recording session. As no significant differences between PRE and POST latencies were found, Wave V latencies from each individual were averaged and compared to normative values that most closely matched our age ranges, intensity level and repetition rate (Jiang et al., 2009). These values are shown in Table 1. A wave V exclusion criteria of Mean Latency +3 Standard Deviations was established to screen for infants that might have generally delayed auditory responses. One infant from each group was excluded from the study due to Wave V latency beyond the cutoff value, giving a final subject pool of 14 infants in each group.

### 3.2. Features of the infant cABR in Quiet and background noise

In general, the waveform morphology of the infant responses in this study appears to be analogous to those observed in immediately older (Johnson et al., 2008) and younger (Anderson et al., 2015) age ranges. Specifically, because the brainstem response to speech so closely reflects the stimulus waveform, three distinct features of the response that mirror the stimulus can be observed (Fig. 1C). The first large positive peak (Wave V), signals the response to speech onset and is similar to the click-evoked Wave V. However, the latency of Wave V to speech is typically later than that to a click, due to the fact that the speech syllable onset has less high-frequency information than the click (Gorga et al., 1989; Stapells et al., 1995). The cABR Wave V, is immediately followed by a negative-going trough, labeled Wave A. In children and adults, a large negativity, peak C, occurs between 12 and 20 ms, signaling the transition period (Russo et al., 2004; Musacchia et al., 2007, 2008; Johnson et al., 2008), but this peak appears to be absent, or under development, in these infant populations. As in previous reports, waves D, E, and F, with a period of ~10 ms between them, define the sustained FFR portion. This portion reflects the phase-locking mechanism of the auditory system in which neurons tend to fire action potentials on one phase of acoustic waveforms at <1000 Hz of stimulation (Marsh et al., 1975; Smith et al., 1975).

### 3.3. Peak detectability in Quiet and background noise

Rater detectability gives a general view of how successful peak-picking can be in this age range. In the YI population, Waves V, A, D and E were detectable in all but one subject and F was detectable in all subjects (Table 2). In OIs, Waves V, A, E and F were also detectable in all but one and, D was picked in all subjects. Peak detectability decreased for both age groups in the Noise condition. In YIs, Peaks V and A were still detectable in all but one subject, but in 11 out of 14 for Waves D and E and in 12 out of 14 for Wave F. In the older group, detectability of Waves V and A dropped to ~70% (10 out of 14), but detectability of FFR peaks remained high.

### 3.4. Relationships between measures of the speech-evoked brainstem response

Pearson's correlations were used to explore the statistical relationships among brainstem measures in the Quiet condition. Tables 3 and 4 show the relationships within onset and across onset and sustained measures, respectively.

Within peak latency measures (Table 3), YIs showed only one relationship: between the positive-going Wave V latency and its

**Table 1**
Click-evoked Response Wave V Latencies and Normative Data (Jiang et al., 2009).

| | Study Data | | | Normative Data | | |
|---|---|---|---|---|---|---|
| Age | Mean (ms) | Std. Dev. (ms) | Min-Max (ms) | Mean (ms) | Std. Dev. (ms) | Norm. Cutoff |
| YI | 6.61 | 0.34 | 6.08–7.03 | 6.36 | 0.24 | 7.12 |
| OI | 6.55 | 0.13 | 6.35–6.80 | 6.12 | 0.25 | 6.87 |

YI: Younger Infants (7–12 months), OI: Older Infants (18–24 months), Std. Dev.: Standard Deviation; ms: milliseconds; Min-Max: range from minimum to maximum values; Norm. Cutoff: Cutoff derived from Normative Data (Mean + 3 SD).

**Table 2**
Transient peak latency measures in infants in quiet and noise conditions.

| Age | Cond. | | Wave V | Wave A | Wave D | Wave E | Wave F |
|---|---|---|---|---|---|---|---|
| YI | Quiet | Mean | 7.72$^{\dagger}$ | 8.95$^{\dagger}$ | 22.61$^{\dagger}$ | 32.19$^{\dagger}$ | 40.46$^{\dagger}$ |
| | | SD | 0.75 | 0.76 | 1.80 | 1.16 | 1.20 |
| | | N | 13 | 13 | 13 | 13 | 14 |
| | Noise | Mean | **8.51** | **10.25** | **23.67** | **34.97** | **42.75** |
| | | SD | 13 | 13 | 11 | 11 | 12 |
| | | N | 1.18 | 1.42 | 2.18 | 1.23 | 1.07 |
| OI | Quiet | Mean | 7.63$^{\dagger}$ | 8.98$^{\dagger}$ | 23.49$^{\dagger}$ | 32.74$^{\dagger}$ | 41.09$^{\dagger}$ |
| | | N | 13 | 13 | 14 | 13 | 13 |
| | | SD | 0.42 | 0.85 | 1.28 | 1.36 | 1.82 |
| | Noise | Mean | **9.47** | **11.10** | **25.62** | **34.50** | **43.14** |
| | | N | 10 | 13 | 14 | 14 | 14 |
| | | SD | 0.98 | 1.11 | 2.16 | 1.78 | 0.95 |

YI: Younger Infants (7–12 months); OI: Older Infants (18–24 months); Cond.: Condition; $^{\dagger}$Main Effect of Noise p-value < 0.05 (Quiet vs. Noise), Longer latencies in bold, SD: Standard Deviation; N: number of subjects.

**Table 3**
Pearson's correlations within Da Quiet transient measures.

| | Younger Infants | | | | Older Infants | | | |
|---|---|---|---|---|---|---|---|---|
| | Wave A | Wave D | Wave E | Wave F | Wave A | Wave D | Wave E | Wave F |
| Wave V | 0.852** | −0.435 | 0.404 | 0.367 | 0.846** | 0.157 | 0.329 | 0.398 |
| Wave A | | −0.248 | 0.336 | 0.267 | | 0.107 | 0.575* | 0.560* |
| Wave D | | | −0.122 | 0.300 | | | −0.269 | −0.348 |
| Wave E | | | | 0.252 | | | | 0.779** |

*$p < 0.05$; **$p < 0.001$; Within-group correlations.

**Table 4**
Pearson's correlations across Da Quiet transient and sustained measures.

| | Younger Infants | | | | | Older Infants | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Wave V | Wave A | Wave D | Wave E | Wave F | Wave V | Wave A | Wave D | Wave E | Wave F |
| F0 Max | 0.246 | 0.129 | −0.525 | −0.138 | −0.182 | 0.007 | 0.024 | −0.110 | −0.363 | −0.060 |
| F0 RMS | 0.426 | 0.305 | −0.553* | −0.192 | −0.087 | −0.030 | 0.025 | −0.263 | −0.286 | 0.069 |
| HF Max | 0.319 | 0.191 | −0.649* | 0.315 | 0.062 | −0.139 | −0.208 | −0.030 | −0.234 | −0.516 |
| HF RMS | 0.446 | 0.291 | −0.667* | 0.457 | 0.132 | −0.271 | −0.364 | −0.073 | −0.377 | −0.612* |
| S-R lag | −0.543* | −0.402 | 0.243 | −0.385 | −0.313 | 0.034 | 0.065 | −0.154 | 0.134 | 0.172 |
| S-R corr | −0.063 | −0.034 | 0.044 | −0.279 | −0.182 | −0.358 | −0.569* | −0.069 | −0.154 | −0.244 |

*$p < 0.05$; **$p < 0.001$; Within-group correlations; F0 Max: Maximum F0 amplitude, F0 RMS: Root Mean Square over the F0 range, HF Max: Maximum Amplitude of the High Frequencies, HF RMS: Root Mean Square over the HF range, S-R lag: Stimulus-to-Response Correlation lag, S-R corr: Stimulus-to-Response Correlation coefficient.

negative-going trough Wave A. OIs also showed the V/A relationship but in addition showed relationships between Wave A, E and F as well as between Waves E and F.

Across transient and sustained measures (Table 4), the YIs had several relationships, particularly between the latency of Wave D and FFR magnitude measures. OIs showed fewer strong and significant relationships.

### 3.5. Effects of age and noise on transient peak latencies in infant responses to speech

Table 2 shows the means, standard deviations and results of the 2X2 ANOVA tests of Condition (Quiet, Noise) X Age (YI, OI) for waves V-F latencies. Main Effects of Noise were observed for Waves V ($F_{1,20} = 28.051$, $p < 0.001$), A ($F_{1,20} = 21.785$, $p < 0.001$), D ($F_{1,21} = 35.165$, $p < 0.001$), E ($F_{1,21} = 30.974$, $p < 0.001$) and F ($F_{1,23} = 40.664$, $p < 0.001$). Mean values show that peak latencies were later in Noise in both OI and YI groups (Table 2). This suggests that the addition of noise delays the timing of the cABR peak responses to a similar degree in both OI and YI infants. It is important to note that no effect of Age was observed in Quiet or Noise conditions for any peak latencies.

### 3.6. Effects of age and noise on FFR measures of infant responses to speech

Fig. 2 shows Grand Average responses in Quiet and Noise conditions for both Younger and Older Infants. Fig. 3 shows grand average spectral amplitudes (computed by FFT) in Quiet and Noise for Younger (Fig. 3A) and Older Infants (Fig. 3B). Table 5 reports the means, standard deviations for the 2X2 ANOVA tests of Condition (Quiet, Noise) X Age (YI, OI) on the FFR measures. Interaction Effects were observed at F0, both in Maximum Amplitude ($F_{1,26} = 4.808$, $p = 0.037$) and RMS ($F_{1,26} = 7.550$, $p = 0.011$) mea-

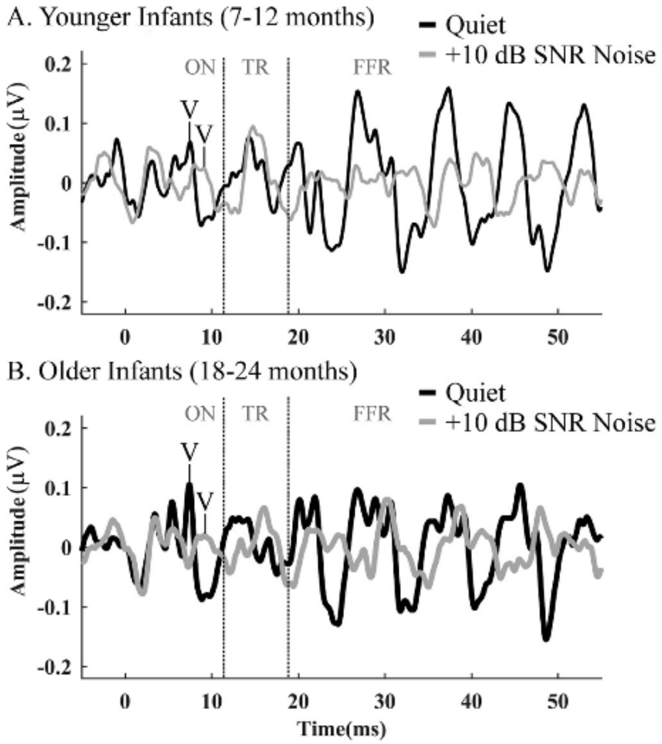

**Fig. 2.** Grand average brainstem responses to /da/ in Quiet and [da]+10 dB SNR. (A) Response features are separated into Onset (ON), Transition (TR) and Frequency Following sections accordingly. The response in Noise follows the general morphology of the Quiet response in YI (7–12 months), but the magnitude and latency in the FFR is degraded. (B) The response in Noise follows the general morphology of the Quiet response in OI (18–24 months) as well, with the temporal features of the waveform in Noise following the features of the Quiet response very closely.

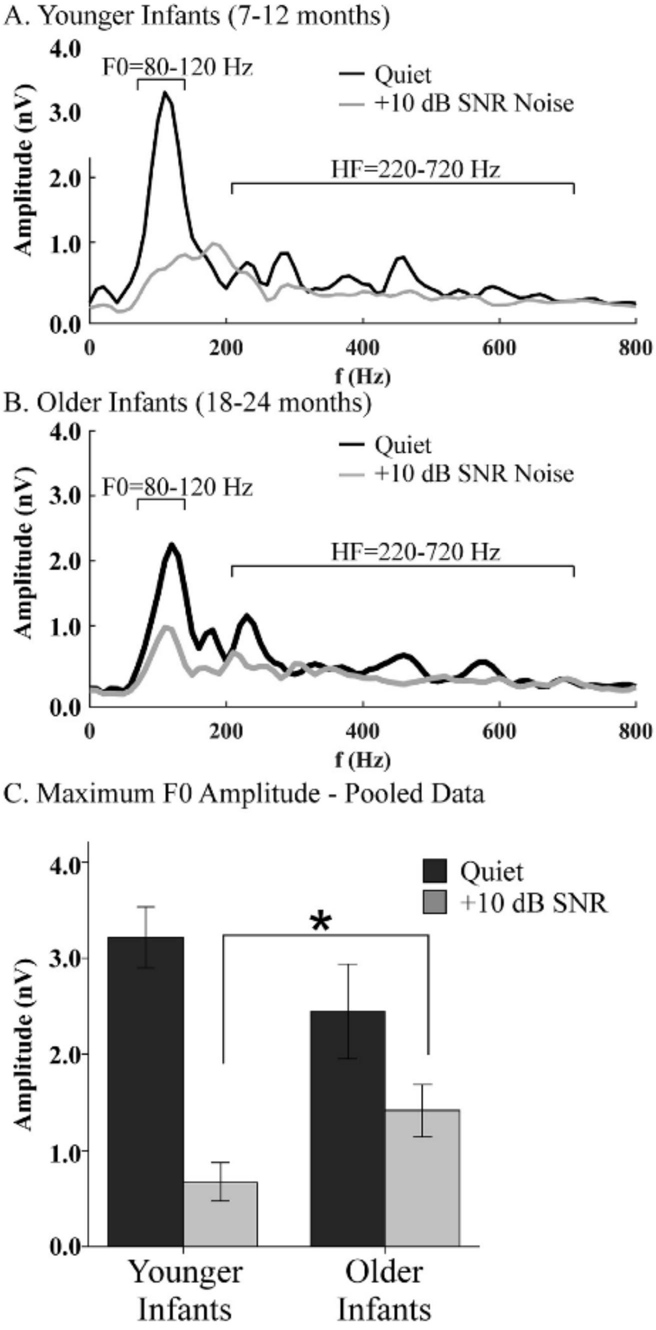

**Fig. 3.** Larger F0 amplitude in Noise in OI infants (18–24 months), compared to YI (7–12 months). (A) Grand average Fast Fourier Transform analysis of the Frequency Following Response (FFR) period (25–40 ms) for YI shows a distinct peak at the stimulus fundamental frequency (F0 = ∼100 Hz) in Quiet (black line), but a diminished magnitude in Noise (grey line). Representation of the higher frequencies (HF) is measured over 220–720 Hz, which accords to the first formant (F1) in the stimulus. (B) The same conditions are shown for OI. The Noise response more closely follows the F0 representation in Quiet in this group, compared to YIs. (C) Pooled data show the maximum F0 peak amplitude in Noise (grey bars) is significantly larger in OIs, compared to YIs.

surements. Main Effects of Noise were also observed in these two tests ($p < 0.001$). Subsequent independent t-tests showed that F0 magnitude in Noise was greater in the OI than in the YI group for both Maximum Amplitude ($t_{26} = 2.190$, $p = 0.038$) and Root Mean Square ($t_{26} = 2.135$, $p = 0.042$) over a range of 80–120 Hz (Fig. 3C). In addition, paired samples test showed that only the YI group had reduced F0 amplitude in Noise as measured by Maximum Amplitude ($t_{13} = 6.501$, $p < 0.001$) or RMS ($t_{13} = 6.522$, $p < 0.001$). OI showed no statistical difference between F0 in Quiet and F0 in Noise ($p > 0.05$).

Similar to the F0 finding, the fidelity of the FFR in noise, measured by Stimulus-to-Response (SR) cross-correlations, also appears to be more resilient in older infants. The SR cross-correlation is a calculation of similarity between the stimulus and response (see Methods) and produces values of lag (time difference between stimulus and response) and r-value (strength of similarity). A Main Effect of Age was observed for the SR r-values ($F_{1,26} = 5.632$, $p = 0.025$). Subsequent t-tests showed that OI r-values were higher in the Noise condition, compared to YIs ($t = 2.524$, $p = 0.018$) (for means see Table 5, for illustration see Fig. 4A and B). No differences in quiet were observed in SR lag times across YI (Mean = 8.22 ms, SD = 0.956 ms) and OI (Mean = 8.40, SD = 0.974)

In contrast to the group differences shown above, HF encoding appears to be impacted by noise similarly in both infant groups. Only a Main Effect of Noise was observed for HF magnitude ($F_{1,26} = 4.347$, $p = 0.047$), and no differences of group were observed in the ensuing t-tests (Table 5).

## 4. Discussion

To the authors' knowledge, this study is the first to show speech-in-noise processing in infants. Based on previous data in children and adults, we predicted that noise disruption of the F0 and HF speech frequencies would be greater at younger ages. We show that older infant encoding of speech F0 amplitude is more robust, and less vulnerable to the addition of +10 dB SNR noise, compared to younger infants. The Stimulus-to-Response (S-R) correlation finding corroborates the F0 amplitude finding, suggesting that the older group has better representation of the periodic portion of speech in noise.

Overall, the data show that older infants (18–24 months) have stronger F0 representation of speech-in-noise compared to their younger counterparts (7–12 months). However, the brainstem response suggests that HF representation is not yet robust enough in either group to reveal differences in within-subject or between-group comparisons. We believe this may be due to a constraint of the brainstem response measurement tool, rather than a reflection of infant capacity for representation of HF at this age, as it is well known that infants can hear sounds above 10 kHz in the first year of life (Trehub et al., 1989).

Previous data in school-aged children showed a clear response pattern (Russo et al., 2004) demonstrating that: (1) transient measures of peak latency are correlated to each other, and (2) sustained FFR measures are correlated to each other, but (3) transient measures and sustained measures are largely independent. The larger number of correlations observed among transient measures in the OI group mirrors the effects reported in school-aged children.

Russo et al. (2004) also showed that noise typically delays the wave V onset response and decreases the magnitude of an individual's FFR response in Quiet by about 30–40% in children and adults. The noise-induced wave V onset delay observed in both groups is in line with the amount of delay seen in children and young adults. This suggests that the impact of noise on broadband, transient acoustic shifts has matured by 7 months of age. In contrast, we found that the magnitude of F0 decrease was larger than in school-aged children or adults; about a 60% decrease in the OI group and up to an 80% decrease in the YI (see Table 5) and these values significantly differed across groups. The Stimulus-to-Response correlation measures exhibited similar results with OIs having a higher SR correlation in Noise, compared to YIs. High frequency representation and peak latency values were similarly

**Table 5**
Measures of the Frequency Following Response in infants for quiet and noise conditions.

| Age | Cond. | | F0 Max. (nV) | F0 RMS (nV) | HF RMS (nV) | S-R corr. (r) |
|---|---|---|---|---|---|---|
| YI | Quiet | M | 3.21 | 2.49 | 0.417 | 0.133 |
| n = 14 | | SD | 1.19 | 0.90 | 0.135 | 0.047 |
| | Noise | M | 0.672 | 0.653 | 0.245 | 0.131 |
| | | SD | 0.757 | 0.451 | 0.075 | 0.052 |
| OI | Quiet | M | 2.45 | 1.71 | 0.295 | 0.148 |
| n = 14 | | SD | 1.86 | 1.26 | 0.120 | 0.058 |
| | Noise | M | 1.42 | 1.10 | 0.249 | 0.184 |
| | | SD | 1.02 | 0.714 | 0.122 | 0.060 |

YI: Younger Infants (7–12 months); OI: Older Infants (18–24 months); M: Mean; SD: Standard Deviation; F0 Max: Maximum F0 amplitude, F0 RMS: Root Mean Square over the F0 range, HF Max: Maximum Amplitude of the High Frequencies, HF RMS: Root Mean Square over the HF range; S-R corr.: Stimulus-to-Response Correlation coefficient; nV: nanovolts.

impacted by noise in both groups, and no group differences were observed for these values in Quiet or Noise conditions.

In line with the Anderson et al. (2015) findings, our data suggest that F0, HF encoding and peak latency maturation of the brainstem response to speech in Quiet may be complete or at the very least, changing quite slowly, by 6-months-of age. Longer time windows between groups and comparison with adult values are needed to determine the entire trajectory of maturation in Quiet. On the other hand, speech processing in background noise appears to have a longer maturational timeline, such that changes continue to occur in speech-related auditory mechanisms well into the second year of life. The results of this study also demonstrate that this developmental timeline can be captured by measuring the F0 magnitude of the FFR response or the Stimulus-to-Response correlations in at least +10 dB SNR background noise.

The current study's data support a maturation-timeline hypothesis, showing that neuronal synchrony of the ABR and FFR peaks is more cohesive in the older group. Peak timing in the current data was evaluated at Waves V, A, D, E and F. In school aged-children and adults, the first two waves are generally related to each other, and are thought to reflect rapid encoding of the frequency shifts in the consonant onset. Waves D, E and F form another group of related waves, reflecting the mechanism of sustained phase-locking to the vowel F0. While both infant groups in this study show a positive correlation between onset Waves V and A, only the older infant group shows a positive correlation between Waves D, E and F latency (Table 3). This implies that the older groups' phase-locking, a measure that reflects phase synchrony of the neural response, is more consistent, such that latency increase in one peak is matched at the next.

Cohesive neuronal synchrony is critical for accurate encoding in noisy conditions because the addition of background sounds randomly interferes with neuronal temporal variability. In the mature mammalian system, single-unit firing of the auditory nerve and brainstem structures such as the cochlear nucleus and inferior colliculus can be time-locked to a segment of the cycle of periodic stimulus (Rose et al., 1967), producing a tight distribution of discharges around a "preferred phase". The width of spike distribution around the preferred phase reflects response variability plus background activity (such as the oscillating patterns of spontaneous firing activity). When background noise is added, the firing distribution pattern widens because noise introduces randomly timed excitatory activity. The interaction between noise and neuronal firing gives rise to a systemic model that helps explain our findings here (see Fig. 5). If neuronal cohesion and synchrony is strong the random firing pattern introduced by the addition of noise will have a small effect. However, if the peripheral system is still immature or a neuronal dys-synchrony is present very early on, the random firing pattern of additional noise will decimate the tenuous frequency following response. In the first case, the neuronal representation of the stimulus will withstand the effect of noise and provide the auditory system with a higher chance of perceptual salience. This appears to be the case with the OI, who have little reduction in F0 in the Noise condition. In contrast, YIs have a nearly decimated F0 response in the Noise condition implying perhaps, weaker phase coherence. In this case there is less chance that the sound will be heard and understood in noise because the major acoustic features related to F0 are indistinct.

Formants, such as F0, are important speech cues that reflect frequency bands of concentrated energy created by the shape of the vocal tract during speech production. F0 provides the perceptual cues needed to discriminate prosodic cues that communicate emotion in speech and can help segregate speech from background noise (Stevens and Klatt, 1974; Ganong, 1980; Qin and Oxenham, 2003). Therefore, difficulty encoding the F0, particularly in noisy environments like the NICU, could contribute to decreased speech intelligibility during critical developmental periods, with possible consequences ranging from a negative impact on the construction of the emerging phonetic map to a decreased recognition of the mother's voice (deRegnier et al., 2002). Another developmental effect may be that younger infants are more likely to respond to auditory change, rather than stable acoustic qualities. This suggests that differential attention in younger infants may be allocated more to detecting dynamic acoustic changes rather than processing static cues such as the F0, making it more difficult to identify speech in noise.

The question naturally arises as to why speech-in-noise would be more robustly encoded in older infants, or conversely, so reduced in the younger cohort? We hypothesize that robust encoding in noise is indicative of a more well-developed sound encoding system. Early maturation of the FFR has been demonstrated as young as two (Van Dyke et al., 2017) and three months of age (Jeng et al., 2010). These studies show that measures of speech F0 and F1 harmonics *in quiet* closely resemble and are statistically indistinguishable from adult responses. Maturation of the FFR in quiet, however, does not reflect the entirety of maturation in the auditory nervous system. Aspects of auditory development are known to continue well into the first decade of life, as measured by both neuronal architecture and perceptual behavior (Werner and Gray, 1998). Embryological and immunochemical studies in post-mortem infant brain tissue show that infants between six and 12 months of age have a fully-developed cochlear sensory apparatus [for review, see (Hall, 2000)] and mature brainstem neuronal architecture (Moore, 2002; Moore and Linthicum, 2007). However, Moore and colleagues also demonstrated that only at about 2 years of age does the neuronal organization of the auditory cortex begin to resemble that of adults, with functioning thalamocortical connections and a mature organization of the granular layer that receives these projections. Better encoding in noise in the older infants may well be linked to this cortical development. A prominent theory of subcortical plasticity argues that "the brainstem operates as part of an integrated network of subcortical and

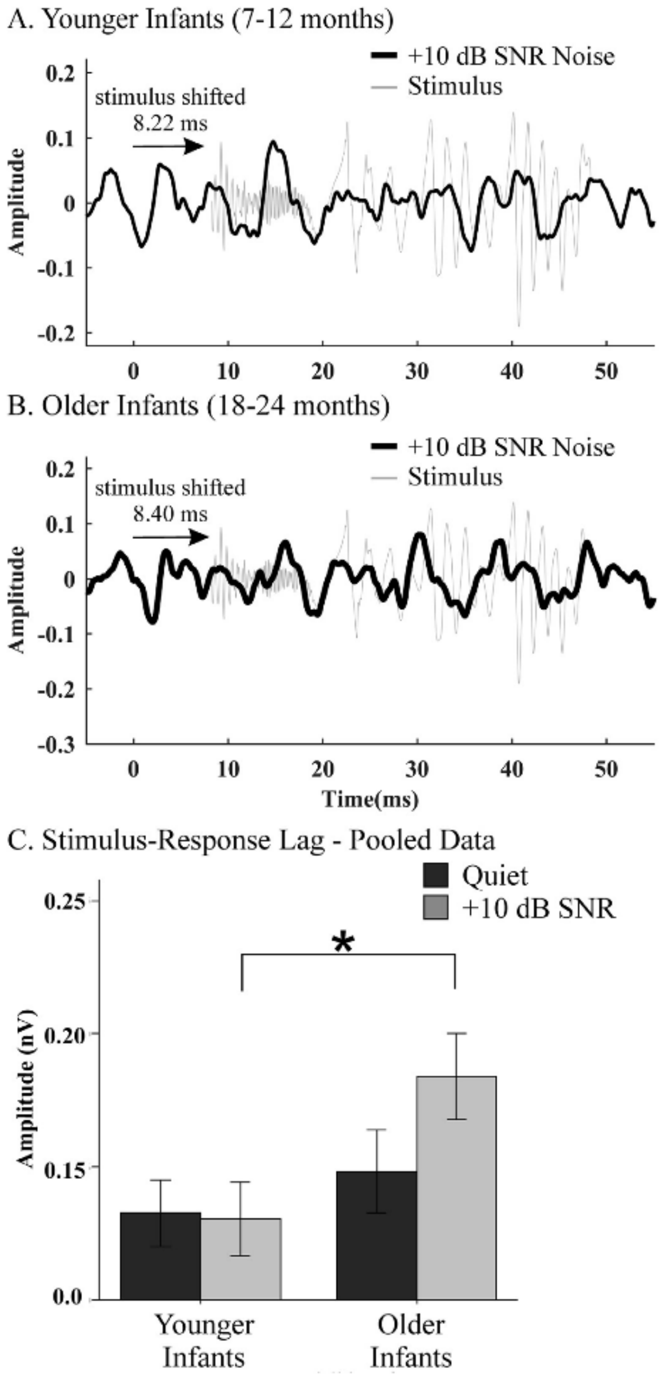

**Fig. 4.** Illustration of Stimulus-to-Response Correlations in Noise and Pooled Means in Quiet and Noise Conditions. (A) The effects of background noise on timing of response were quantified by performing stimulus-to-response (S-R) correlations. Analysis was performed by shifting the Frequency Following Response portion of the response (thick black line) over the entire stimulus until a maximum correlation was achieved. This time period represents the neuronal conduction time or time it takes the stimulus to reach from the ear to the brain. In panel A, the stimulus waveform (thin grey line) is shifted to accord with the mean correlation lag (8.22 ms) for the YI group (7–10 months). Waves D, E and F of the YI response in Noise (thick black line) are shifted, and unevenly timed in relation to the F0 peaks in the stimulus. (B) The same conditions are shown for the OI (18–24 month) group. The stimulus waveform is shifted by the mean lag (8.40 ms). (C) Pooled data for the correlation coefficients are shown with significant differences between groups in the noise condition.

cortical structures linked by afferent and efferent processes" (Song et al., 2012). This theory is supported by animal models showing that plasticity in the brainstem is gated by cortical activity (Hyde and Knudsen, 2001; Suga et al., 2002). If true, this implies that neuronal resistance to the deleterious effects of noise may, in part, depend on cortical activity in the developing system and that infant encoding in noise will be unstable until some threshold of maturation has been reached. A dynamic neuromaturation theory of speech processing reflecting these principles is illustrated in Fig. 5.

From the developmental point of view, phonemic maps of an infant's native language are still under development within auditory cortex between 6 and 12 months-of-age (Ortiz-Mantilla et al., 2013, 2016). By 18–24 months, these phonemic cortical maps are already in place, making infants more efficient language processors. Previous studies posit that the demand for complex organization, such as cortical mapping, and the near-simultaneous encoding of acoustic detail engages cortical mechanisms that are capable of refining the neural code at a basic sensory level (Musacchia et al., 2008; Patel, 2014). This idea is consistent with models of perceptual learning that involve perceptual weighting with feedback (Nosofsky, 1986) and animal models showing recurrent corticofugal modulation during ecologically salient tasks (Liberman and Mattingly, 1989; Zhang et al., 1997; Hyde and Knudsen, 2001, 2002). In infants, establishment of native language phonemic maps in the auditory cortex may increase perceptual weighting and decrease the cognitive effort required to process speech sounds; making way for fine-tuning of encoding mechanisms in brainstem nuclei.

## 5. Conclusions

Overall, this study demonstrates that speech-in-noise processing measures can be recorded in awake, unsedated infants and that the developmental timeline for maturation of speech-in-noise processing extends into the second year of life. These findings underscore the need for an expanded cadre of audiological testing measures in order to obtain a complete picture of functional hearing in infants and young children. The results detailed here expand on prior studies that used meaningful measures of speech and speech-in-noise processing in school-aged children to delineate and quantify the complex auditory processing mechanisms that support language acquisition. Measures of speech and speech-in-noise processing that have been especially useful in identifying neural deficits associated with language problems in older children show great promise as a tool to enable very early identification of currently undetectable deficits in infants' auditory processing.

In order for the cABR to be useful in identifying early markers of auditory deficits, normative values must be established and compared to language outcomes. The demographic and case history data we obtained on our subjects was sparse but targeted (i.e. no known neurological disorders, no birth complications and passed the universal newborn hearing screening). However, we cannot definitively state that no child in our cohort will develop language learning impairments later in life. Therefore, it is premature to suggest that these data can be used as normative values. Furthermore, the present data cannot distinguish whether the absolute amplitude measures of speech-in-noise at either age is more predictive of language outcome, as compared to the degree of change over time. In order to disentangle these questions, follow-up studies are needed that compare early response patterns to later patterns, and language outc0omes in the same infants. In addition, larger samples are needed to account for more cross-age variability in this population. Further exploration of the predictive value of the cABR to later language outcome, given the fact that it appears to be both sensitive and reliable, may well increase the potential utility of the speech-in-noise cABR/FFR measure.

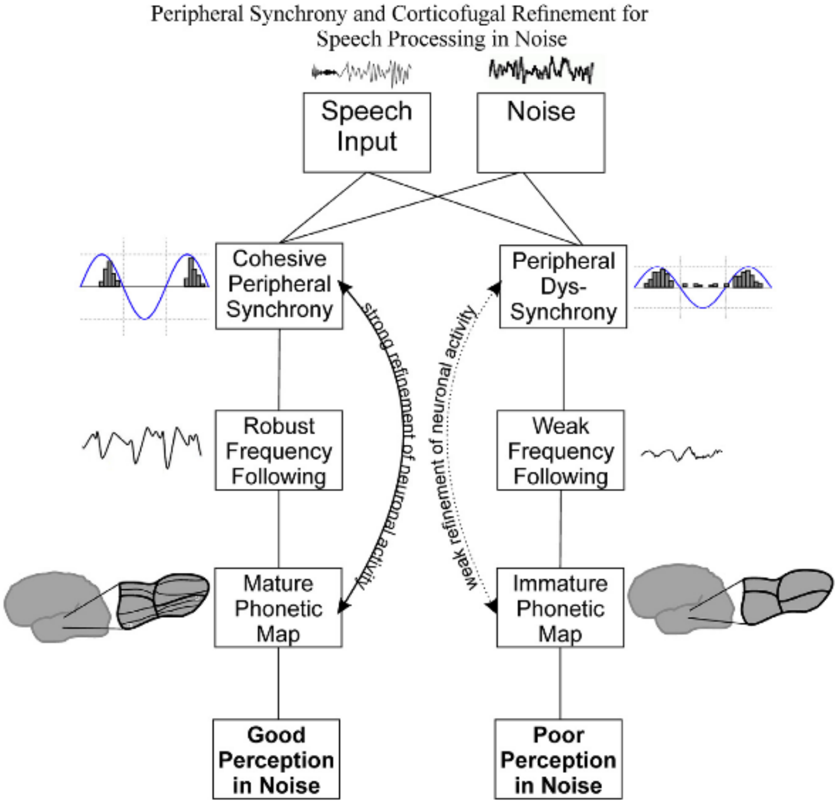

**Fig. 5.** Block diagram to explain the working hypothesis for the perception of speech in noise according to dynamic peripheral and central auditory processing. A simplified model of the auditory system is presented. Speech-in-noise is first transduced from sound vibrations into electrical energy in the peripheral system. If the synchrony of the peripheral auditory system is cohesive, as illustrated by a tightly grouped phase-locking histogram (Left), there will be a robust frequency following response observed in the cABR or FFR scalp-recorded response. A less mature system with a high degree of neuronal dys-synchrony on the other hand, as illustrated by a flat distribution of the phase-locked response on the right, will be reflected in a weak scalp-recorded FFR response. The robustness of the peripheral representation may then determine the degree to which phonetic boundaries are developed, and the extent to which corticofugal modulation of firing takes place. When both the peripheral and central mechanisms are mature, this leads to good perception regardless of noise (left). If the auditory system is immature, or abnormally dys-synchronous, this can lead to a recursive weakening of speech perception in noise.

## Conflict of interest

The authors certify that they have NO affiliations with or involvement in any organization or entity with any financial interest, or non-financial interest (such as personal or professional relationships, affiliations, knowledge or beliefs) in the subject matter or materials discussed in this manuscript.

## Acknowledgements

We want to thank all the families who participated in this study.

## References

Abrams DA, Nicol T, Zecker SG, Kraus N. Auditory brainstem timing predicts cerebral asymmetry for speech. J Neurosci 2006;26:11131–7.

Allen P, Wightman F. Effects of signal and masker uncertainty on children's detection. J Speech Hear Res 1995;38:503–11.

Anderson S, Parbery-Clark A, White-Schwoch T, Kraus N. Aging affects neural precision of speech encoding. J Neurosci 2012;32:14156–64. https://doi.org/10.1523/JNEUROSCI.2176-12.2012.

Anderson S, Parbery-Clark A, White-Schwoch T, Kraus N. Auditory brainstem response to complex sounds predicts self-reported speech-in-noise performance. J Speech Lang Hear Res 2013;56:31–43. https://doi.org/10.1044/1092-4388(2012/12-0043).

Anderson S, Parbery-Clark A, White-Schwoch T, Kraus N. Development of subcortical speech representation in human infants. J Acoust Soc Am 2015;137:3346–55. https://doi.org/10.1121/1.4921032.

Baumann O, Belin P. Perceptual scaling of voice identity: common dimensions for different vowels and speakers. Psychol Res 2010;74:110–20. https://doi.org/10.1007/s00426-008-0185-z.

Benasich AA, Choudhury N, Friedman JT, Realpe-Bonilla T, Chojnowska C, Gou Z. The infant as a prelinguistic model for language learning impairments: predicting from event-related potentials to behavior. Neuropsychologia 2006;44:396–411.

Bhagat S. The frequency following response. In: Atcherson SR, Stoody TM, editors. Auditory electrophysiology: a clinical guide. New York (NY): Thieme; 2012. p. 86–102.

Bradlow AR, Kraus N, Hayes E. Speaking clearly for children with learning disabilities: sentence perception in noise. J Speech Lang Hear Res 2003;46:80–97.

Brandt J, Rosen JJ. Auditory phonemic perception in dyslexia: categorical identification and discrimination of stop consonants. Brain Lang 1980;9:324–37.

Brons I, Houben R, Dreschler WA. Effects of noise reduction on speech intelligibility, perceived listening effort, and personal preference in hearing-impaired listeners. Trends Hear 2014;18. https://doi.org/10.1177/2331216514553924.

Choudhury N, Benasich AA. Maturation of auditory evoked potentials from 6 to 48 months: prediction to 3 and 4 year language and cognitive abilities. Clin Neurophysiol 2011;122:320–38.

Cooper RP, Aslin RN. The language environment of the young infant: implications for early perceptual development. Can J Psychol 1989;43:247–65.

deRegnier R-A, Wewerka S, Georgieff M, Mattia F, Nelson C. Influences of posconceptional age and postnatal experience on the development of auditory recognition memory in the newborn infant. Dev Psychobiol 2002;41:216–25.

Eilers RE. Infant speech perception. Auditory development in infancy. Switzerland: Springer; 1985. p. 197–213. 978-1-4757-9340-6.

Eilers RE, Morse PA, Gavin WJ, Oller DK. Discrimination of voice onset time in infancy. J Acoust Soc Am 1981;70:955–65.

Eisenberg LS. Current state of knowledge: speech recognition and production in children with hearing impairment. Ear Hear 2007;28:766–72. https://doi.org/10.1097/AUD.0b013e318157f01f.

Eggermont JJ. Physiology of the developing auditory system. Auditory development in infancy. Switzerland: Springer; 1985. p. 21–45. 978-1-4757-9340-6.

Faulkner A, Rosen S. Contributions of temporal encodings of voicing, voicelessness, fundamental frequency, and amplitude variation to audio-visual and auditory speech perception. J Acoust Soc Am 1999;106:2063–73.

Galbraith GC, Amaya EM, de Rivera JM, Donan NM, Duong MT, Hsu JN, Tran K, Tsang LP. Brain stem evoked response to forward and reversed speech in humans. Neuroreport 2004;15:2057–60.

Galbraith GC, Arbagey PW, Branski R, Comerci N, Rector PM. Intelligible speech encoded in the human brain stem frequency-following response. Neuroreport 1995;6:2363–7.

Galbraith GC, Jhaveri SP, Kuo J. Speech-evoked brainstem frequency-following responses during verbal transformations due to word repetition. Electroencephalogr Clin Neurophysiol 1997;102:46–53.

Ganong 3rd WF. Phonetic categorization in auditory word perception. J Exp Psychol Human 1980;6:110–25.

Gorga MP, Kaminski JR, Beauchaine KL, Jesteadt W, Neely ST. Auditory brainstem responses from children three months to three years of age: normal patterns of response. II. J Speech Hear Res 1989;32:281–8.

Hall 3rd JW. Development of the ear and hearing. J Perinatol 2000;20:S12–20.

Hall 3rd JW, Rupp KA. Auditory brainstem response: recent developments in recording and analysis. Adv Otorhinolaryngol 1997;53:21–45.

Hall II JW. Handbook of auditory evoked responses needham heights. MA: Allyn and Bacon; 1992.

Helfer KS, Freyman RL. Aging and speech-on-speech masking. Ear Hear 2008;29:87–98. https://doi.org/10.1097/AUD.0b013e31815d638b.

Hind SE, Haines-Bazrafshan R, Benton CL, Brassington W, Towle B, Moore DR. Prevalence of clinical referrals having hearing thresholds within normal limits. Int J Audiol 2011;50:708–16. https://doi.org/10.3109/14992027.2011.582049.

Hoormann J, Falkenstein M, Hohnsbein J, Blanke L. The human frequency-following response (FFR): normal variability and relation to the click-evoked brainstem response. Hear Res 1992;59:179–88.

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

Hornickel J, Skoe E, Nicol T, Zecker S, Kraus N. Subcortical differentiation of stop consonants relates to reading and speech-in-noise perception. Proc Natl Acad Sci USA 2009;106:13022–7.

Horwitz AR, Dubno JR, Ahlstrom JB. Recognition of low-pass-filtered consonants in noise with normal and impaired high-frequency hearing. J Acoust Soc Am 2002;111:409–16.

Hyde PS, Knudsen EI. A topographic instructive signal guides the adjustment of the auditory space map in the optic tectum. J Neurosci 2001;21:8586–93.

Hyde PS, Knudsen EI. The optic tectum controls visually guided adaptive plasticity in the owl's auditory space map. Nature 2002;415:73–6.

Jamieson DG, Kranjc G, Yu K, Hodgetts WE. Speech intelligibility of young school-aged children in the presence of real-life classroom noise. J Am Acad Audiol 2004;15:508–17.

Jeng FC, Schnabel EA, Dickman BM, Hu J, Li X, Lin CD, Chung HK. Early maturation of frequency-following responses to voice pitch in infants with normal hearing. Percept Mot Skills 2010;111:765–84. https://doi.org/10.2466/10.22.24.PMS.111.6.765-784.

Jiang ZD, Wu YY, Wilkinson AR. Age-related changes in BAER at different click rates from neonates to adults. Acta Paediatr 2009;98:1284–7. https://doi.org/10.1111/j.1651-2227.2009.01312.x.

Johnson KL, Nicol T, Zecker SG, Kraus N. Developmental plasticity in the human auditory brainstem. J Neurosci 2008;28:4000–7.

Johnson KL, Nicol TG, Kraus N. Brain stem response to speech: a biological marker of auditory processing. Ear Hear 2005;26:424–34.

Keogh T, Kei J, Driscoll C, Khan A. Children with minimal conductive hearing impairment: speech comprehension in noise. Audiol Neurootol 2010;15:27–35. https://doi.org/10.1159/000218360.

King C, Warrier CM, Hayes E, Kraus N. Deficits in auditory brainstem pathway encoding of speech sounds in children with learning problems. Neurosci Lett 2002;319:111–5.

Klatt DH. Software for a cascade/parallel formant synthesizer. J Acoust Soc Am 1980;67:971–95.

Kraus N, Nicol T. Brainstem origins for cortical 'what' and 'where' pathways in the auditory system. Trends Neurosci 2005;28:176–81.

Kraus NA S, White-Schoch T, Fay RR, Popper AN. The frequency-following response a window into human communication. Springer; 2017.

Krishnan A, Xu Y, Gandour J, Cariani P. Encoding of pitch in the human brainstem is sensitive to language experience. Cogn Brain Res 2005;25:161–8.

Laukli E, Mair IW. Low-frequency sensorineural hearing loss. Brainstem response, speech and high-frequency audiometry. Scand Audiol 1985;14:133–9.

Leibold LJ, Buss E. Children's identification of consonants in a speech-shaped noise or a two-talker masker. J Speech Lang Hear Res 2013;56:1144–55. https://doi.org/10.1044/1092-4388(2012/12-0011).

Leibold LJ, Neff DL. Masking by a remote-frequency noise band in children and adults. Ear Hear 2011;32:663–6. https://doi.org/10.1097/AUD.0b013e31820e5074.

Leibold LJ, Yarnell Bonino A, Buss E. Masked Speech Perception Thresholds in Infants, Children, and Adults. Ear Hear 2016;37:345–53. https://doi.org/10.1097/AUD.0000000000000270.

Lenneberg EH. Biological foundations of language. New York: John Wiley and Sons; 1967.

Liberman AM, Mattingly IG. A specialization for speech perception. Science 1989;243:489–94.

Lieu JE. Speech-language and educational consequences of unilateral hearing loss in children. Arch Otolaryngol Head Neck Surg 2004;130:524–30. https://doi.org/10.1001/archotol.130.5.524.

Luinge MR, Post WJ, Wit HP, Goorhuis-Brouwer SM. The ordering of milestones in language development for children from 1 to 6 years of age. J Speech Lang Hear Res 2006;49:923–40. https://doi.org/10.1044/1092-4388(2006/067).

Marsh JT, Brown WS, Smith JC. Far-field recorded frequency-following responses: correlates of low pitch auditory perception in humans. Clin Neurophysiol 1975;38:113–9.

McCreery RW, Spratford M, Kirby B, Brennan M. Individual differences in language and working memory affect children's speech recognition in noise. Int J Audiol 2017;56:306–15. https://doi.org/10.1080/14992027.2016.1266703.

Miller JL, Jusczyk PW. Seeking the neurobiological bases of speech perception. Cognition 1989;33:111–37.

Millman RE, Mattys SL. Auditory verbal working memory as a predictor of speech perception in modulated maskers in listeners with normal hearing. J Speech Lang Hear Res 2017;60:1236–45. https://doi.org/10.1044/2017_JSLHR-S-16-0105.

Moeller MP, Tomblin JB. An introduction to the outcomes of children with hearing loss study. Ear Hear 2015;36(Suppl. 1):4S–13S. https://doi.org/10.1097/AUD.0000000000000210.

Moore DR, Edmondson-Jones M, Dawes P, Fortnum H, McCormack A, Pierzycki RH, Munro KJ. Relation between speech-in-noise threshold, hearing loss and cognition from 40–69 years of age. PLoS One 2014;9:e107720. https://doi.org/10.1371/journal.pone.0107720.

Moore JK. Maturation of human auditory cortex: implications for speech perception. Ann Otol Rhinol Laryngol Suppl 2002;189:7–10.

Moore JK, Linthicum Jr FH. The human auditory system: a timeline of development. Int J Audiol 2007;46:460–78. https://doi.org/10.1080/14992020701383019.

Muir DW. The development of infants' auditory spatial sensivity. Auditory development in infancy. Switzerland: Springer; 1985. p. 51–83. 978-1-4757-9340-6.

Musacchia G, Sams M, Skoe E, Kraus N. Musicians have enhanced subcortical auditory and audiovisual processing of speech and music. Proc Natl Acad Sci USA 2007;104:15894–8.

Musacchia G, Strait D, Kraus N. Relationships between behavior, brainstem and cortical encoding of seen and heard speech in musicians and non-musicians. Hear Res 2008;241:34–42.

Nittrouer S, Tarr E, Wucinich T, Moberly AC, Lowenstein JH. Measuring the effects of spectral smearing and enhancement on speech recognition in noise for adults and children. J Acoust Soc Am 2015;137:2004–14. https://doi.org/10.1121/1.4916203.

Nosofsky RM. Attention, similarity, and the identification-categorization relationship. J Exp Psychol Gen 1986;115:39–61.

Oh EL, Wightman F, Lutfi RA. Children's detection of pure-tone signals with random multitone maskers. J Acoust Soc Am 2001;109:2888–95.

Ortiz-Mantilla S, Hämäläinen JA, Musacchia G, Benasich AA. Enhancement of gamma oscillations indicates preferential processing of native over foreign phonemic contrasts in infants. J Neurosci 2013;33:18746–54.

Ortiz-Mantilla S, Hämäläinen JA, Realpe-Bonilla T, Benasich AA. Oscillatory dynamics underlying perceptual narrowing of native phoneme mapping from 6 to 12 months of age. J Neurosci 2016;36:12095–105. https://doi.org/10.1523/JNEUROSCI.1162-16.2016.

Patel AD. Can nonlinguistic musical training change the way the brain processes speech? The expanded OPERA hypothesis. Hear Res 2014;308:98–108. https://doi.org/10.1016/j.heares.2013.08.011.

Qin MK, Oxenham AJ. Effects of simulated cochlear-implant processing on speech reception in fluctuating maskers. J Acoust Soc Am 2003;114:446–54.

Rose JE, Brugge JF, Anderson DJ, Hind JE. Phase-locked response to low-frequency tones in single auditory nerve fibers of the squirrel monkey. J Neurophysiol 1967;30:769–93.

Rubel EW, Ryals BM. Development of the place principle: acoustic trauma. Science 1983;219:512–4.

Russo N, Nicol T, Musacchia G, Kraus N. Brainstem responses to speech syllables. Clin Neurophysiol 2004;115:2021–30.

Skoe E, Kraus N. Auditory brain stem response to complex sounds: a tutorial. Ear Hear 2010;31:302–24.

Skoe E, Krizman J, Anderson S, Kraus N. Stability and plasticity of auditory brainstem function across the lifespan. Cereb Cortex 2015;25:1415–26. https://doi.org/10.1093/cercor/bht311.

Smith DR, Patterson RD. The interaction of glottal-pulse rate and vocal-tract length in judgements of speaker size, sex, and age. J Acoust Soc Am 2005;118:3177–86.

Smith JC, Marsh JT, Brown WS. Far-field recorded frequency-following responses: evidence for the locus of brainstem sources. Electroencephalogr Clin Neurophysiol 1975;39:465–72.

Soderlund GB, Jobs EN. Differences in speech recognition between children with attention deficits and typically developed children disappear when exposed to 65 dB of auditory noise. Front Psychol 2016;7:34. https://doi.org/10.3389/fpsyg.2016.00034.

Sohmer H, Pratt H, Kinarti R. Sources of frequency following responses (FFR) in man. Electroencephalogr Clin Neurophysiol 1977;42:656–64.

Song JH, Nicol T, Kraus N. Test-retest reliability of the speech-evoked auditory brainstem response. Clin Neurophysiol 2011a;122:346–55. https://doi.org/10.1016/j.clinph.2010.07.009.

Song JH, Skoe E, Banai K, Kraus N. Perception of speech in noise: neural correlates. J Cogn Neurosci 2011b;23:2268–79. https://doi.org/10.1162/jocn.2010.21556.

Song JH, Skoe E, Banai K, Kraus N. Training to improve hearing speech in noise: biological mechanisms. Cereb Cortex 2012;22:1180–90. https://doi.org/10.1093/cercor/bhr196.

Sperling AJ, Lu ZL, Manis FR, Seidenberg MS. Deficits in perceptual noise exclusion in developmental dyslexia. Nat Neurosci 2005;8:862–3. https://doi.org/10.1038/nn1474.

Spitzer SM, Liss JM, Mattys SL. Acoustic cues to lexical segmentation: a study of resynthesized speech. J Acoust Soc Am 2007;122:3678–87. https://doi.org/10.1121/1.2801545.

Stapells DR, Gravel JS, Martin BA. Thresholds for auditory brain stem responses to tones in notched noise from infants and young children with normal hearing or sensorineural hearing loss. Ear Hear 1995;16:361–71.

Stelmachowicz PG, Pittman AL, Hoover BM, Lewis DE, Moeller MP. The importance of high-frequency audibility in the speech and language development of children with hearing loss. Arch Otolaryngol Head Neck Surg 2004;130:556–62. https://doi.org/10.1001/archotol.130.5.556.

Stevens KN, Klatt DH. Role of formant transitions in the voiced-voiceless distinction for stops. J Acoust Soc Am 1974;55:653–9.

Suga N, Xiao Z, Ma X, Ji W. Plasticity and corticofugal modulation for hearing in adult animals. Neuron 2002;36:9–18.

Trehub SE, Schneider BA, Morrongiello BA, Thorpe LA. Developmental changes in high-frequency sensitivity. Audiology 1989;28:241–9.

Van Dyke KB, Lieberman R, Presacco A, Anderson S. Development of phase locking and frequency representation in the infant frequency-following response. J Speech Lang Hear Res 2017:1–12. https://doi.org/10.1044/2017_JSLHR-H-16-0263.

Vance M, Martindale N. Assessing speech perception in children with language difficulties: effects of background noise and phonetic contrast. Int J Speech Lang Pathol 2012;14:48–58. https://doi.org/10.3109/17549507.2011.616602.

Werker JF, Gilbert JH, Humphrey K, Tees RC. Developmental aspects of cross-language speech perception. Child Dev 1981;52:349–55.

Werker JF, Hensch TK. Critical periods in speech perception: new directions. Annu Rev Psychol 2015;66:173–96.

Werker JF, Tees RC. Developmental changes across childhood in the perception of non-native speech sounds. Can J Psychol 1983;37:278–86.

Werker JF, Tees RC. Speech perception as a window for understanding plasticity and commitment in language systems of the brain. Dev Psychobiol 2005;46:233–51.

Werner LA, Gray L. Behavioral studies of hearing development. In: Rubel EW, Popper AN, Fay RR, editors. Development of the auditory system. New York (NY): Springer; 1998.

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

Zhang Y, Suga N, Yan J. Corticofugal modulation of frequency processing in bat auditory system. Nature 1997;387:900–3.

Ziegler JC, Pech-Georgel C, George F, Lorenzi C. Noise on, voicing off: Speech perception deficits in children with specific language impairment. J Exp Child Psychol 2011;110:362–72. https://doi.org/10.1016/j.jecp.2011.05.001.