

# Exponential Modeling of Frequency-Following Responses in American Neonates and Adults

DOI: 10.3766/jaaa.16135

Fuh-Cherng Jeng\*

Brandie Nance\*

Karen Montgomery-Reagan†

Chia-Der Lin‡

## Abstract

**Background:** The scalp-recorded frequency-following response (FFR) has been widely accepted in assessing the brain's processing of speech stimuli for people who speak tonal and nontonal languages. Characteristics of scalp-recorded FFRs with increasing number of sweeps have been delineated through the use of an exponential curve-fitting model in Chinese adults; however, characteristics of speech processing for people who speak a nontonal language remain unclear.

**Purpose:** This study had two specific aims. The first was to examine the characteristics of speech processing in neonates and adults who speak a nontonal language, to evaluate the goodness of fit of an exponential model on neonatal and adult FFRs, and to determine the differences, if any, between the two groups of participants. The second aim was to assess effective recording parameters for American neonates and adults.

**Research Design:** This investigation employed a prospective between-subject study design.

**Study Sample:** A total of 12 American neonates (1–3 days old) and 12 American adults ( $24.1 \pm 2.5$  yr old) were recruited. Each neonate passed an automated hearing screening at birth and all adult participants had normal hearing and were native English speakers.

**Data Collection and Analysis:** The English vowel /i/ with a rising pitch contour (117–166 Hz) was used to elicit the FFR. A total of 8,000 accepted sweeps were recorded from each participant. Three objective indices (Frequency Error, Tracking Accuracy, and Pitch Strength) were computed to estimate the frequency-tracking acuity and neural phase-locking magnitude when progressively more sweeps were included in the averaged waveform. For each objective index, the FFR trends were fit to an exponential curve-fitting model that included estimates of asymptotic amplitude, noise amplitude, and a time constant.

**Results:** Significant differences were observed between groups for Frequency Error, Tracking Accuracy, and Pitch Strength of the FFR trends. The adult participants had significantly smaller Frequency Error ( $p < 0.001$ ), better Tracking Accuracy ( $p = 0.001$ ), and larger Pitch Strength ( $p = 0.003$ ) values than the neonate participants. The adult participants also demonstrated a faster rate of improvement (i.e., a smaller time constant) in all three objective indices compared to the neonate participants. The smaller time constants observed in adults indicate that a larger number of sweeps will be needed to adequately assess the FFR for neonates. Furthermore, the exponential curve-fitting model provided a good fit to the FFR trends with increasing number of sweeps for American neonates (mean  $r^2 = 0.89$ ) and adults (mean  $r^2 = 0.96$ ).

**Conclusions:** Significant differences were noted between the neonatal and adult participants for Frequency Error, Tracking Accuracy, and Pitch Strength. These differences have important clinical implications in determining when to stop a recording and the number of sweeps needed to adequately

\*Communication Sciences and Disorders, Ohio University, Athens, OH; †OhioHealth O'Bleness Hospital, Athens, OH; ‡Department of Otolaryngology-HNS, China Medical University Hospital, Taiwan, China

Corresponding author: Fuh-Cherng Jeng, Division of Communication Sciences and Disorders, Ohio University, Athens, OH 45701; E-mail: jeng@ohio.edu

This study was supported by National Science Foundation, Division of Behavioral and Cognitive Sciences (grant number BCS-1250700) and Ohio University – Baker Fund Award (grant number BA-15-07).

assess the frequency-encoding acuity and neural phase-locking magnitude in neonates and adults. These findings lay an important foundation for establishing a normative database for American neonates and adults, and may prove to be useful in the development of diagnostic and therapeutic paradigms for neonates and adults who speak a nontonal language.

**Key Words:** auditory-evoked potentials, exponential model, frequency-following response, neonate

**Abbreviations:** F0 = fundamental frequency; FFR = frequency-following response; RECD = real-ear-to-coupler difference; SNR = signal-to-noise ratio

## INTRODUCTION

The scalp-recorded frequency-following response (FFR) has shown its potential to help us better understand the speech processing mechanisms and neural plasticity of the human brain. For example, it has been reported that the FFR provides an objective method to understand speech processing mechanisms for typically developing children and children with deficient auditory function (Russo et al, 2004). Some children with autism spectrum disorders have shown less frequency-tracking acuity than typically developing children (Russo et al, 2008). It has also been reported that with short-term training on specific linguistic pitch contours, listeners not only improve their behavioral response accuracy but also express enhanced frequency-tracking acuity (Song et al, 2008). Hornickel et al (2009) reported that the FFR elicited by speech sounds was related to reading and speech-in-noise processing for 8- to 13-yr-old school-age children.

The brain's processing of speech stimuli, as reflected by the scalp-recorded FFR, has been reported for normal-hearing adults who speak a tonal (Krishnan et al, 2005; Swaminathan et al, 2008) or nontonal (Galbraith et al, 2004; Aiken and Picton, 2006; 2008) language. The accurate encoding of voice pitch is also important for listeners to process and appreciate music. Recent studies (Musacchia et al, 2007; Wong et al, 2007; Skoe and Kraus, 2013) have shown that musical training enhances the acuity of frequency tracking in the human brain. Intracranial electrophysiological studies have shown that the FFR can also be observed at the cortical level, including the primary auditory cortex (Behroozmand et al, 2016) and beyond (Nourski et al, 2013). These findings support the notion that the FFR is a viable and objective neurophysiological index of the brain's processing of speech sounds. Most importantly, these findings also have potential clinical applications for diagnostic and remediation strategies for normal and pathological populations.

One major challenge in recording the FFR is its relatively low signal-to-noise ratio (SNR). The FFR is a small electrophysiological response with an amplitude that is usually in the range of 0.1–0.3  $\mu\text{V}$ , whereas the background noise (environmental or physiological) is much larger in amplitude with a typical range of  $\sim 10$ – $20$   $\mu\text{V}$  (Jeng, Chung, et al, 2011). In order to examine the

frequency-encoding acuity and neural phase-locking magnitude for neonates and adults, the SNR of an FFR recording must be improved. For example, with a limited number of sweeps (e.g., 100 sweeps or less), the response is usually not detectable. On the other hand, with abundant sweeps (e.g., 8,000 sweeps), the response likely will be robust and detectable. However, recording 8,000 sweeps will take a much longer time than recording 100 sweeps. Thus, the question becomes: What is the least number of sweeps that will be needed to obtain a reliable response? One possible solution is to collect data for up to, for example, 8,000 sweeps. Based on the data of 8,000 sweeps, a computer model can be derived to fit the increasing trend of the response. When a computer model has been developed for neonates, a set of response-threshold criteria can be derived to determine the least number of sweeps that can be used to signal the presence of a response and subsequently shorten the amount of time needed to complete a recording for neonates in future applications. The same principle also applies when developing a computer model for adults.

In a previous study, Jeng, Chung, et al (2011) have shown a general trend regarding the robustness of an FFR as a function of the number of recorded sweeps in Chinese adults. A preliminary computer model has been successfully applied to capture the response trend and to derive a set of response-threshold criteria that can be used to determine if an FFR is present for Chinese adults with normal hearing. However, neural circuitry and functional organization of the brain in adults who speak a nontonal language are likely different from those in Chinese adults. Additionally, whether such a computational model would provide a good fit to neonates during their immediate postnatal days remains unknown. Thus, a computer model with specific parameters that are precisely tailored for each population is a critical and much-needed component to shorten the testing time needed for neonates and adults who speak a nontonal language. Development of such a computer model is a necessary step to help us better understand the nature of speech processing inside the brain and to derive useful diagnostic and therapeutic protocols down the line.

The primary purpose of this study was to develop a computer model capable of capturing FFR trends as a function of sweeps for neonates and adults who speak a nontonal language. Further, this study had two

specific aims. The first was to examine the dependency of the neonatal and adult FFRs on the number of recorded sweeps and to evaluate any differences in the resultant modeling coefficients such as the time constants and asymptotic amplitudes between the two groups of participants. The second aim was to assess effective recording parameters for American neonates and adults. Given the success of using such a computer model in an adult population who speaks a tonal language (Jeng, Chung, et al, 2011), it was hypothesized that such a computer model would still be feasible in delineating the FFR trends for people who speak a non-tonal language. Additionally, because a neonate's auditory system is relatively immature when compared to that of adults (Rubel and Ryals, 1983), it was hypothesized that the American adults would exhibit a stronger response trend (i.e., faster improvements and larger response amplitudes in speech encoding with increasing number of sweeps) than neonates. In other words, the number of sweeps needed to satisfy a certain response-threshold criterion would be larger for neonates than adults.

## MATERIALS AND METHODS

### Participants

Twelve American neonates (6 girls and 6 boys, age 1–3 days) were recruited from the OhioHealth O'Bleness Hospital in Athens, OH. All neonates were both healthy and full-term, born to native speakers of American English. These neonates were negative of any syndromic or neurologic disorders and were free from any known risk factors of hearing impairment (Joint Committee on Infant Hearing, 1994) such as low birth weight, hyperbilirubinemia, and low Apgar scores. Each neonate passed an automated auditory brainstem hearing screening (ALGO 3i; Natus Medical Incorporated, San Carlos, CA) before the completion of the following experiments.

To determine the difference of speech representation between immediate postnatal days and adulthood, 12 American adults (11 females and 1 male, age  $24.1 \pm 2.5$  yr) were also recruited. These adult participants were native speakers of American English. All the adult participants had normal hearing, as determined by pure-tone thresholds  $\leq 20$  dB HL for octave frequencies between 250 and 8000 Hz. None of the adult participants had received  $>3$  yr of musical training and thus were considered as nonmusicians.

Hearing screenings and FFR recordings for the neonates were conducted in a quiet room in the Newborn Center at the OhioHealth O'Bleness Hospital; recordings of the adult participants were conducted in an acoustically treated sound booth at Ohio University. During the experiments, neonates were fast asleep or

in a state of rest. Adult participants were seated in a comfortable recliner with their eyes closed and were encouraged to relax and fall asleep. All experimental protocols and data analysis procedures were approved by the institutional review board (IRB number: 15X093) at Ohio University.

### Stimulus Presentation

The English vowel /i/ with a rising pitch contour (117–166 Hz) was used to elicit the FFR because this stimulus is commonly used in FFR literature (Krishnan et al, 2004; 2005; Aiken and Picton, 2006; 2008) and had been previously used in our publications (Jeng et al, 2010; 2013; Jeng, Chung, et al, 2011, Jeng, Hu, 2011). This stimulus had a duration of 250 msec with a 10-msec rise and fall time of the stimulus envelope. Stimulus presentation and trigger synchronization were controlled by custom-built software written in LabVIEW (National Instruments, Austin, TX). The silent interval between the offset of a stimulus and the onset of the next stimulus was set at 45 msec, which produced a stimulation rate of 3.39 stimuli/sec. The stimulus was presented to each participant for up to a total of 8,800 times, producing a test time of  $\sim 43$  min ( $295 \text{ msec} \times 8,800 \text{ sweeps} = \sim 43$  min) for each participant. The stimulus was delivered monaurally via an electromagnetically shielded insert earphone (Etymotic [Elk Grove Village, IL] ER-3A) to the right ear at 75 dB SPL for the adult participants.

Considering the relatively smaller ear canal volumes that are commonly observed in neonates, the published age-appropriate real-ear-to-coupler difference (RECD) values were applied to control for the sound pressure level differences inside the neonate and adult's ear canals. Specifically, the RECD value for 1-mo-old infants at 250 Hz is  $\sim 5$  dB larger than that measured in adults (Feigin et al, 1989; Keefe et al, 1993; Scollie et al, 1998; Dillon, 2001). To compensate for the difference of sound pressure levels inside the neonate and adult's ear canals, the stimulus was presented at 70 dB SPL for neonate participants. The RECD value at 250 Hz for 1-mo-old infants was used because a value at 125 Hz for neonates was unavailable.

### Recording Parameters

Recording parameters were identical for neonates and adults. Specifically, three gold-plated recording electrodes were mounted to the high forehead along the midline below the hairline (noninverting), right mastoid (inverting), and low forehead (ground). All electrode impedances were kept below  $3,000 \Omega$  at 10 Hz. Continuous brain waves were amplified through an Intelligent Hearing Systems (Miami, FL) OptiAmp amplifier with a gain of 100,000, band-pass filtering at 10–3000 Hz, and a slope of 6 dB/octave. The continuous

brain waves were routed through a 16-bit analog-to-digital converter and were sampled at a rate of 20,000 samples/sec by using a National Instruments input-output control card (model number USB-6216 BNC). All recordings were saved on a computer for off-line analysis.

## Data Analysis

All recordings were analyzed using bespoke scripts written in MATLAB (MathWorks, Natick, MA). The data analysis procedures were similar to those reported in our previous publications (Jeng, Chung, et al, 2011; Jeng et al, 2013). Briefly, each recording was digitally band-pass filtered through a brick wall, linear phase finite-impulse-response filter. Each filtered recording was then segmented into sweeps of 295 msec in length. An individual sweep was rejected if it contained voltages greater than  $\pm 25 \mu\text{V}$ . For each participant, the rejection rate was  $< 10\%$ , and a total of 8,000 artifact-free sweeps were included in the averaging procedure. Recordings obtained from a distinct number of sweeps, starting from the first sweep, were included in the averaged recordings. The number of sweeps used in the averaging procedure were 1, 10, 20, 50, 100, 200, 500, 800, 1,000, 1,200, 1,400, 1,600, 1,800, 2,000, 2,200, 2,400, 2,600, 2,800, 3,000, 3,500, 4,000, 5,000, 6,000, 7,000, and 8,000. Each averaged recording was subject to the following analysis procedures. First, the time waveform of each averaged recording and the stimulus time waveform went through a cross-correlation function. After that, the time shift that produced the maximum cross-correlation value between 3 and 10 msec after the onset of the stimulus was identified (Galbraith et al, 2001; Russo et al, 2005). Then, starting from the time shift that produced the maximum cross-correlation value, a 250-msec segment of the recording was extracted from the averaged waveform. And finally, the same procedure was applied to each averaged waveform of all recordings obtained from the neonate and adult participants.

## Extraction of Fundamental Frequency Contours

A sliding-window, narrow-band spectrogram was used to extract the spectral information of each averaged recording. Specifically, a 50-msec Hanning window was applied to each averaged recording with a step size of 1 msec, which resulted in a total of 201 time bins to be analyzed. Each time bin was zero-padded to 1 sec and thus provided a 1-Hz frequency resolution in the spectrogram. For each time bin, the frequency that corresponded to the maximal spectral density was searched within a predefined frequency range and determined as the fundamental frequency (F0) estimate for that time bin. This procedure was repeated for all 201 time bins. All F0 estimates were concatenated to constitute the F0 contour of an averaged recording. A predefined frequency range

(107–176 Hz) was used to fit with the F0 contour of the speech stimulus and allow a buffer of 10 Hz for error measurements. The same procedure was applied to each averaged recording and the stimulus waveform.

## Objective Indices

Three objective indices (Frequency Error, Tracking Accuracy, and Pitch Strength) were used to estimate the frequency-encoding acuity and neural phase-locking magnitude for each averaged recording. These objective indices were chosen because each of them represented an important aspect of the brain's processing of speech stimuli. These indices were commonly used in FFR literature (Krishnan et al, 2004; 2005) and in our previous publications (Jeng et al, 2010; Jeng, Chung, et al, 2011; Jeng, Hu, et al, 2011). Definitions and computations of the three indices are described as follows:

1. Frequency Error indicated the accuracy of frequency encoding in response to a speech stimulus. To compute a Frequency Error, the F0 estimates of the speech stimulus were subtracted from those of an averaged recording. The absolute values of these differences between the F0 estimates of the stimulus and an averaged recording were then averaged across the 201 time bins to constitute a Frequency Error for each averaged recording.
2. Tracking Accuracy represented the overall faithfulness of frequency encoding during the time course of stimulus presentation. This index was defined as the linear regression  $r$  value on a recording-versus-stimulus F0 contours plot.
3. Pitch Strength denoted the magnitude of neural phase locking that included the amplitude at F0 and its harmonics. This index was computed by measuring the difference between the positive peaks within 5–10 msec to its following negative trough in a normalized autocorrelation function of each averaged recording.

## Exponential Modeling of the FFR Trends

Measurements of each objective index (i.e., Frequency Error, Tracking Accuracy, and Pitch Strength) were subject to an exponential curve-fitting model. For Tracking Accuracy and Pitch Strength, which had ascending trends with increasing number of sweeps, the following model was used to determine the dependency of response trends on the number of sweeps that had been included in the averaging procedure:

$$A(n) = A_{AS} \left( 1 - e^{-n/\tau} \right) - A_{noise} \quad (1)$$

where  $A$  represents an objective index (i.e., Tracking Accuracy or Pitch Strength) of the FFR;  $n$  is the number of sweeps included in the averaging procedure;  $A_{AS}$

denotes the asymptotic amplitude of the response and is computed from the fitted curve of the exponential model with the number of sweeps being 8,000;  $A_{\text{noise}}$  annotates the amplitude of noise and is derived from the fitted curve of the FFR trend when the number of sweeps equals 1;  $e$  is Euler's mathematical constant = 2.7182; and  $\tau$  denotes the time constant of the fitted curve that indicates the number of sweeps needed to reach 63% of the asymptotic amplitude. Computation of the time constant  $\tau$  at its 63% asymptotic amplitude is based on the mathematical theorem that any exponential function with base  $e$  is identical to its derivative (Courant and Robbins, 1996; Goldstein et al, 2009). For example, when the number of sweeps “ $n$ ” included in the averaged waveform reaches the time constant  $\tau$ , the ascending exponential model with zero noise will be  $A(n) = A_{\text{AS}}(1 - e^{-n/\tau}) = A_{\text{AS}}(1 - e^{-1}) = 0.63A_{\text{AS}}$ .

For Frequency Error, which had a descending trend with increasing number of sweeps, an alternative model was used to determine the dependency of the response trend on the number of sweeps included in the averaged waveform. In this case,  $A_{\text{noise}}$  and  $A_{\text{AS}}$  were switched due to the nature of a descending trend:

$$A(n) = A_{\text{noise}} \left( e^{-n/\tau} \right) - A_{\text{AS}} \quad (2)$$

A repeated measures two-way analysis of variance was used to determine the significance of the maturity of the auditory system (i.e., neonates versus adults) and the number of sweeps. A  $p < 0.05$  indicated a statistically significant difference.

## RESULTS

The neonatal and adult FFRs were visualized by plotting the amplitude spectrograms as a function of the number of sweeps included in the averaging procedure. Figure 1 plots the grand-averaged spectrograms for neonates (upper panels) and adults (lower panels) with increasing number of sweeps. For both groups of participants, the FFRs were difficult to distinguish from the background noise when  $\leq 500$  sweeps were included in the averaging procedure. However, the FFR became

more prominent when the number of sweeps was progressively increased up to 8,000 sweeps.

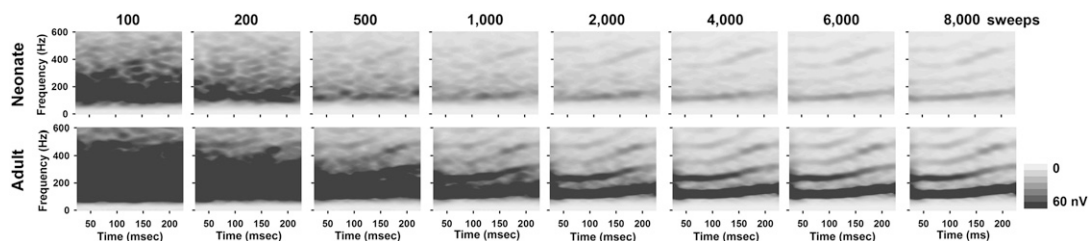
## Quantitative Analyses of the FFR Trends

In order to quantify the FFRs obtained in the two groups of participants, all recordings were processed by using the abovementioned methodology. Figure 2 displays a typical example of the F0 contour of a response (left panel) and the autocorrelation output (right panel) of a recording obtained from an adult participant when 8,000 sweeps were included in the averaging procedure. In terms of the frequency-encoding acuity (left panel), the F0 contour of the response generally followed the F0 contour of the stimulus. In terms of the magnitude of neural phase-locking (right panel), autocorrelation output of the same recording revealed an overall periodicity of the response. Pitch Strength was calculated by measuring the peak-to-trough amplitude starting from the first positive peak to its following negative trough in the normalized autocorrelation output. The response F0 contour and autocorrelation output curve shown in this figure are typical of those observed in both neonates and adults.

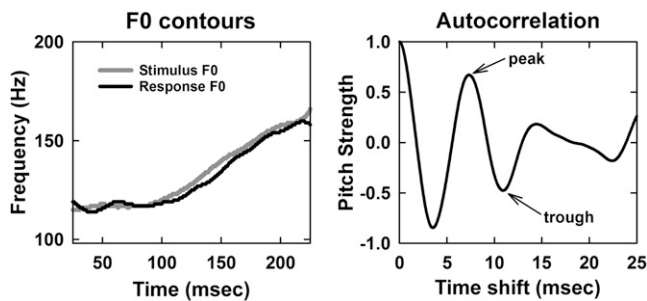
## FFR Trends With Respect To the Different Objective Indices

Three indices were used to quantify the frequency-tracking acuity and neural phase-locking magnitude. Figure 3A–3C plots Frequency Error, Tracking Accuracy, and Pitch Strength, respectively, as a function of number of sweeps for neonates and adults. Recordings obtained from these two groups of participants are plotted in the same panels for comparison.

For both groups of participants, relatively large values of Frequency Errors (Figure 3A) were observed when only a limited number of sweeps, such as 1 or 10 sweeps, were included in the averaging procedure. Frequency Error estimates declined substantially with increasing number of sweeps and appeared to reach



**Figure 1.** Narrow-band spectrograms of the FFRs elicited by the English vowel /i/ with a rising pitch contour in American neonates (1–3 days old; top row) and American adults (bottom row). The numbers on the top of each column indicate the number of sweeps included in the averaging procedure. Spectrograms were derived by using a sliding Hanning window with a duration of 50 msec, step size of 1 msec, and frequency resolution of 1 Hz. A gray scale on the right indicates the amplitudes in nV for the neonatal and adult spectrograms.



**Figure 2.** A typical example of the F0 contour (left panel) and autocorrelation output (right panel) of an FFR elicited by the English vowel /i/ with a rising pitch contour.

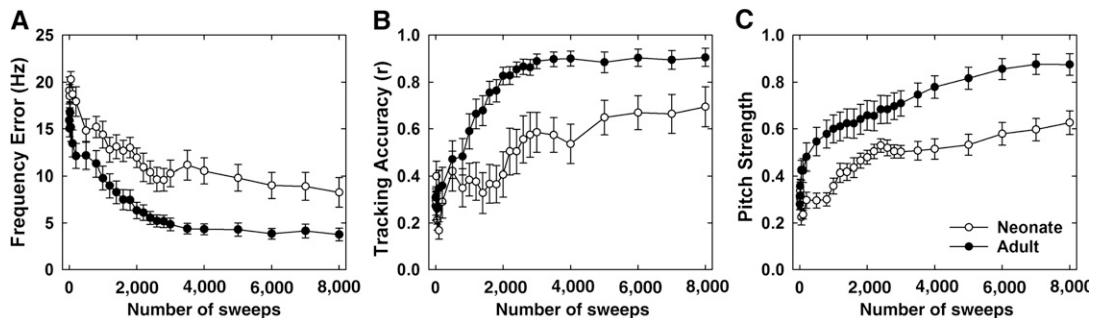
an asymptotic amplitude. Specifically, for adults, Frequency Error estimates were  $\sim 17$  Hz when  $\leq 10$  sweeps were included. The Frequency Error estimates then declined with increasing number of sweeps and reached a steady state of  $\sim 4$  Hz at  $\sim 4,000$ – $8,000$  sweeps. For neonates, Frequency Error estimates showed a similar trend and declined from  $\sim 19$  to 9 Hz. Although the neonates and adults showed similar values of Frequency Errors for low numbers of sweeps, Frequency Errors declined at a faster rate and reached a smaller asymptotic amplitude in adults than those in neonates.

Tracking Accuracy (Figure 3B) showed an increasing trend with increasing number of sweeps. For adults, their Tracking Accuracy estimates were  $\sim 0.26$  at 10 sweeps and increased substantially with increasing number of sweeps, reaching an asymptotic amplitude of  $\sim 0.9$  at  $4,000$ – $8,000$  sweeps. For neonates, their Tracking Accuracy estimates showed a similar increasing trend. However, Tracking Accuracy estimates obtained in the neonate participants increased at a slower rate and reached a smaller asymptotic amplitude than those obtained in the adult participants. Pitch Strength estimates obtained in the two groups of participants (Figure 3C) also showed similar increasing trends as those observed in Tracking Accuracy.

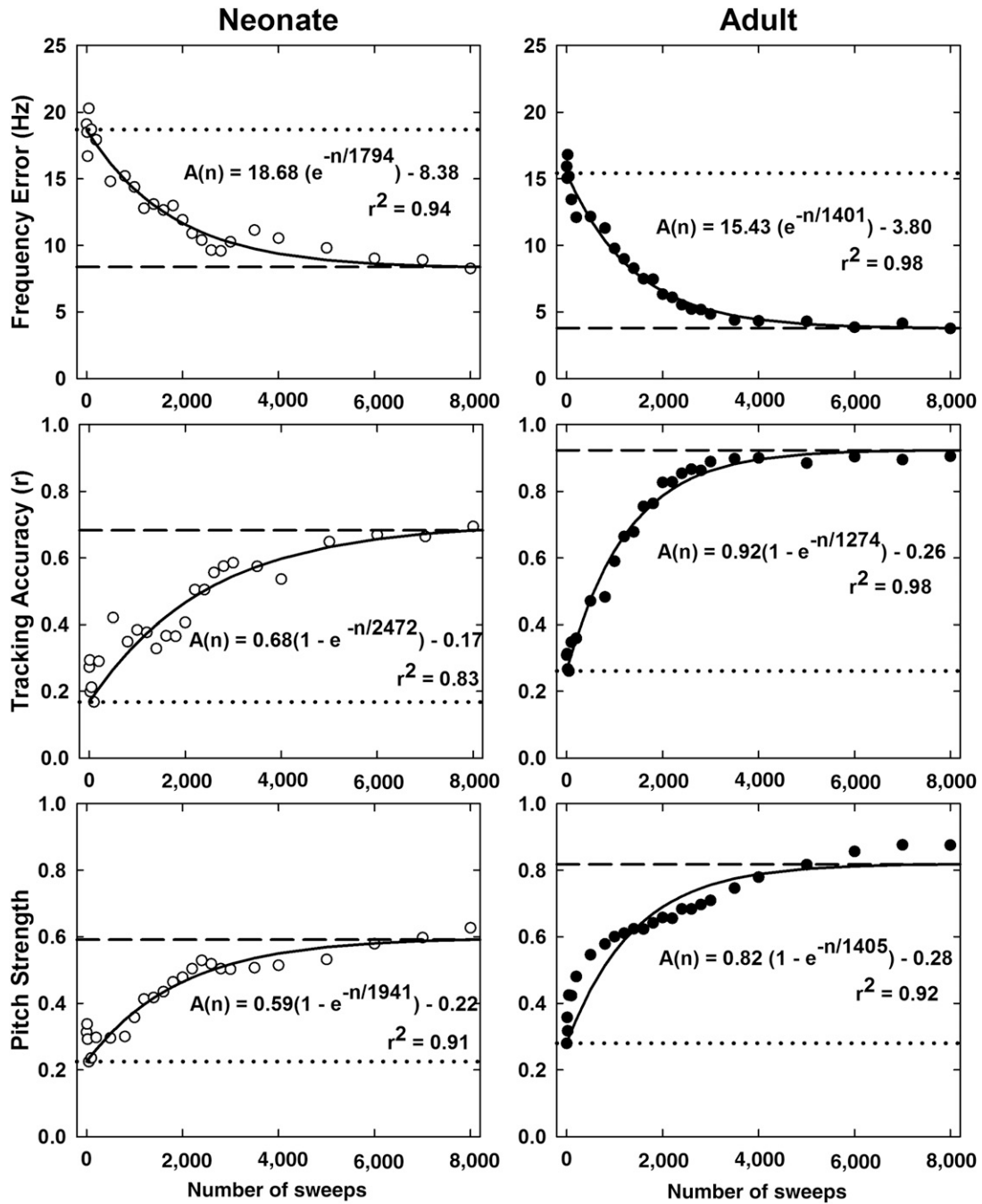
### Exponential Modeling of FFR Trends in Neonates and Adults

To better quantify the decreasing and increasing FFR trends, data obtained in the neonates and adults were fit to the abovementioned descending or ascending exponential model. Figure 4 shows the exponential curves that best fit the response trends of the Frequency Error, Tracking Accuracy, and Pitch Strength estimates obtained in the neonate (left column) and adult (right column) participants. The exponential model provided a good fit to the FFR trends, with a mean  $r^2$  value of 0.96 across the three objective indices for the adult participants and a mean  $r^2$  value of 0.89 across the three objective indices for the neonate participants. The dotted line used in the left and middle columns indicates the noise amplitude ( $A_{noise}$ ) of the fitted curve and is calculated from the fitted curve of the FFR trends when the number of sweeps equals 1, whereas the dashed line used in the left and middle columns indicates the response asymptotic amplitude ( $A_{AS}$ ) of the fitted curve and is calculated from the fitted curve of the FFR trends when the number of sweeps equals 8,000. The resultant exponential equations and coefficients of determination (i.e.,  $r^2$ ) are displayed in each panel.

The asymptotic amplitude ( $A_{AS}$ ) of the Frequency Errors (Figure 4, top row) derived from the fitted curve was 8.38 Hz for the neonate participants and 3.80 Hz for the adult participants. The noise amplitudes ( $A_{noise}$ ) of the Frequency Errors derived from the fitted curves for the neonate and adult participants were 18.68 and 15.43 Hz, respectively. The unfavorable noise amplitudes indicated low SNRs when only a limited number of sweeps was included in the averaging procedure. In addition to the asymptotic and noise amplitudes, another important parameter of the exponential model was the time constant ( $\tau$ ) of the fitted curve, which represented the number of sweeps required to reach 63% of the asymptotic amplitude. The  $\tau$  values of the fitted curves for the Frequency Error FFR trends obtained in the neonate and adult participants were 1,794 and



**Figure 3.** FFR trends for (A) Frequency Error, (B) Tracking Error, and (C) Pitch Strength, with increasing number of sweeps for American neonates (open circles) and adults (filled circles). Data obtained from American neonates and adults are plotted on the same panels for comparison. Vertical bars represent 1 standard error.



**Figure 4.** Exponential modeling of the FFR trends for American neonates (left column) and adults (right column) with respect to the three objective indices: Frequency Error (top row), Tracking Accuracy (middle row), and Pitch Strength (bottom row). The fitted equations, along with their coefficients of determination ( $r^2$ ), are annotated in each panel. The dashed and dotted horizontal lines indicate the estimated asymptotic and noise amplitudes of the FFR recordings, respectively.

1,401 sweeps, respectively. It is important to note that the adult participants produced a smaller  $\tau$  value than neonates. This finding indicated that the adult participants had a faster improvement in frequency-encoding acuity (i.e., less Frequency Error) with increasing number of sweeps compared to the neonates. For Frequency Error, significant differences were observed between the two groups of participants ( $F = 17.313, p < 0.001$ ,

$\eta_p^2 = 0.382$ ) and across the number of sweeps ( $F = 40.871, p < 0.001, \eta_p^2 = 0.593$ ).

Tracking Accuracy (Figure 4, middle row) of the fitted curves demonstrated an ascending trend with increasing number of sweeps. The asymptotic amplitudes of fitted curves for the Tracking Accuracy trends derived from the neonate and adult participants were 0.68 and 0.92, respectively. The noise amplitude of the fitted curve for

Tracking Accuracy was 0.17 for the neonate participants and 0.26 for the adult participants. The  $\tau$  values of the fitted curve for the neonate and adult participants were 2,472 and 1,274 sweeps, respectively. Note the adult participants demonstrated a faster rate of increment (i.e., a smaller  $\tau$  value) and reached a better frequency-tracking acuity estimate (i.e., a larger asymptotic amplitude) than neonates. For Tracking Accuracy, significant differences were also observed between the two groups of participants ( $F = 14.837, p = 0.001, \eta_p^2 = 0.346$ ) and across the number of sweeps ( $F = 28.352, p < 0.001, \eta_p^2 = 0.503$ ).

Pitch Strength (Figure 4, bottom row) showed similar ascending trends compared to those observed in Tracking Accuracy. The asymptotic amplitude of the fitted curve for Pitch Strength was 0.59 for the neonate participants and 0.82 for the adult participants. Noise amplitudes of the fitted curves for the neonate and adult participants were 0.22 and 0.28, respectively. The  $\tau$  value of the fitted curve was 1,941 sweeps for the neonate participants and 1,405 sweeps for the adult participants. The adult participants reached a larger enhancement (i.e., a larger asymptotic amplitude) at a faster rate (i.e., a smaller  $\tau$  value) than neonates. For Pitch Strength, significant differences were observed between the two groups of participants ( $F = 10.605, p = 0.003, \eta_p^2 = 0.275$ ) and across the number of sweeps ( $F = 41.921, p < 0.001, \eta_p^2 = 0.600$ ). For clarity, results of the curve-fitting coefficients (i.e., the asymptotic amplitudes ( $A_{AS}$ ), noise amplitudes ( $A_{noise}$ ), and  $\tau$  values) of the exponential models are summarized in Table 1.

## DISCUSSION

Given the increasing potential of the scalp-recorded FFR in the realms of basic research and clinical applications, frequency-encoding acuity and neural phase-locking magnitude become important factors to examine. This is particularly true for neonates who are too young to provide reliable behavioral results. This study used an exponential curve-fitting model that included up to 8,000 artifact-free sweeps and examined the FFR trends for neonates and adults. Results demonstrated that the exponential model provided a good fit to the FFR trends for the American neonates (mean  $r^2 = 0.89$ ) and adults (mean  $r^2 = 0.96$ ). This finding

indicates the relevance of using such an exponential model to mathematically analyze the FFR trends in neonates during their immediate postnatal days and in adults who speak a nontonal language.

### Effects of Number of Sweeps

Among the many methods that can be used to improve the SNR of a recording, signal averaging is one of the most commonly used methodologies to reduce the amplitude of the background noise. When more sweeps are progressively included in the averaging procedure, physiological responses that are phase-locked to the stimulus presentation will build up with increasing number of sweeps, whereas the amplitude of the background noise (due to its nature of randomness) will approach to a mean of zero. The SNR of an FFR recording thus improves with increasing number of sweeps.

Several mathematical models have been developed to delineate the various response trends of the auditory system (Nourski et al, 2005; Miller et al, 2006; Jeng, Chung, et al, 2011). For example, Miller et al (2006) used a decaying exponential model and successfully delineated the time course of the firing probability of auditory nerve fibers in cats. Nourski et al (2005) used a double exponential curve-fitting algorithm and successfully described the time course of the compound action potentials of the auditory nerve in guinea pigs. For the scalp-recorded FFR, Jeng, Chung, et al (2011) used an exponential model and successfully delineated the FFR trends as a function of number of sweeps in Chinese adults. Results obtained in this present study indicate that such an exponential model provides a good fit to the FFR trends for American neonates during their immediate postnatal days and American adults who speak a nontonal language.

### Exponential Model of the FFR Trends in Neonates and Adults

For the first time, response trends of frequency-encoding acuity and neural phase-locking magnitude were successfully fitted to an exponential model for neonates during their immediate postnatal days. Although neural elements of the auditory system are fully

**Table 1. Exponential Curve-Fitting Coefficients of the FFR Trends Obtained in American Neonates and Adults**

	Frequency Error (Hz)			Tracking Accuracy ( <i>r</i> )			Pitch Strength		
	$A_{AS}$	$A_{noise}$	$\tau$	$A_{AS}$	$A_{noise}$	$\tau$	$A_{AS}$	$A_{noise}$	$\tau$
Neonate	8.38	18.68	1,794	0.68	0.17	2,472	0.59	0.22	1,941
Adult	3.80	15.43	1,401	0.92	0.26	1,274	0.82	0.28	1,405

Note:  $A_{AS}$  = asymptotic amplitude of the fitted curve;  $A_{noise}$  = noise amplitude of the fitted curve;  $\tau$  = time constant of the fitted curve indicating the number of sweeps required to reach 63% of the asymptotic amplitude.



matured during the first few days after birth, results obtained in the present study revealed that neonates 1–3 days old demonstrated readily measurable FFRs with increasing number of sweeps.

When testing neonates and adults, the amount of time needed to complete a recording is an important factor to consider. Data obtained from the exponential curve-fitting of the FFR trends also provide guidance on when to stop a recording. For example, if FFR recordings are conducted in American neonates and Pitch Strength is used to signal the presence of a response, the  $A_{AS}$ ,  $A_{noise}$ , and  $\tau$  estimates derived from the exponential model can be used to estimate the number of sweeps needed to stop a recording. For example, if 80% of the asymptotic amplitude is the targeted amplitude to obtain,  $\sim 3,125$  sweeps ( $1.61 \tau = 1.61 \times 1,941 = 3,125$ ) will need to be included in the averaging procedure for American neonates. However, for American adults to reach 80% of their asymptotic amplitude, on average only 2,262 sweeps ( $1.61 \tau = 1.61 \times 1,405 = 2,262$ ) are required. That is, while 2,262 sweeps will be sufficient for adults, a larger number of sweeps will be needed for neonates to adequately assess the FFR. The same notation can also be applied when Frequency Error or Tracking Accuracy is used to signal the presence of a response.

When enough sweeps were included in the averaging procedure, adult participants demonstrated better frequency-tracking acuity and larger asymptotic amplitude than neonates. Additionally, adult participants demonstrated a faster improvement (i.e., a smaller  $\tau$  value) in all three objective indices (Frequency Error, Tracking Accuracy, and Pitch Strength) with increasing number of sweeps when compared to the neonates. These findings can be attributed, at least partially, to the maturity and development of the auditory system. Effects of the development of the auditory system on the neural processing of speech stimuli have been reported in the FFR literature (Jeng et al, 2010; Jeng, Hu, et al, 2011; Anderson et al, 2015; Skoe et al, 2015). For example, Jeng et al (2010) reported an increment in speech representation through a longitudinal case study of one American infant. FFRs obtained from this American infant showed improvement in speech processing from age 1 mo to 11 mo. Anderson et al (2015) recruited 25 American infants ranging in age from 3 mo to 10 mo and demonstrated a rapid maturational change as a function of age. All these results attest to the theory that, although English is a nontonal language, voice pitch and intonation still carry important cues for speech understanding in English. In addition to the matured auditory system in adults, the improved frequency-encoding acuity and neural phase-locking magnitude for the adult participants can be

attributed, at least partially, to the adult participants' exposure to the various voice pitch and intonation cues that are present in their native language.

Future studies including neonates and adults who speak tonal and nontonal languages may shed light on the relative contributions of the maturation of the auditory system and the listener's linguistic experience on the brain's processing of speech stimuli. For example, native versus nonnative speech stimuli can be used to elicit FFRs in American and Chinese neonates and adults. Differences between the four groups of participants in response to native versus nonnative speech stimuli may indicate the relative contributions of the maturation of the auditory system and the listener's linguistic experience on FFRs. This current study is the beginning of a larger project. It is important to note that, although this present study demonstrates results (i.e., the exponential model provides a good fit to FFR trends in neonates and adults) consistent with those of an earlier study (Jeng, Chung, et al, 2011), data from these two studies were obtained by using different equipment and experimental protocols. Future studies including both American and Chinese neonates and adults using identical equipment and experimental protocols will be needed to allow direct comparisons between the four groups of participants.

## CONCLUSIONS

Significant differences were observed between American neonates and adults for Frequency Error, Tracking Accuracy, and Pitch Strength. Specifically, American adults had significantly better frequency-tracking acuity and larger phase-locking magnitude than neonates during their immediate postnatal days. Additionally, American adults demonstrated a faster improvement rate in both frequency-tracking acuity and neural phase-locking magnitude than American neonates. The faster improvement rate in adults, as reflected by a smaller time constant, indicates that the number of sweeps needed to adequately assess the FFR is larger for neonates than for adults. The exponential curve-fitting model provides a good fit to the FFR trends with increasing number of sweeps for both the American neonates and adults. These findings lay an important foundation in the development of a normative database for neural processing of speech sounds for people who speak a nontonal language.

**Acknowledgments.** The authors thank the neonates and their families and adults who participated in this study. The authors also thank Hallie Ganch, Kathryn Kulasa, and Megan Presley for their assistance in data collection.

## REFERENCES

- Aiken SJ, Picton TW. (2006) Envelope following responses to natural vowels. *Audiol Neurootol* 11(4):213–232.
- Aiken SJ, Picton TW. (2008) Envelope and spectral frequency-following responses to vowel sounds. *Hear Res* 245(1–2):35–47.
- Anderson S, Parbery-Clark A, White-Schwoch T, Kraus N. (2015) Development of subcortical speech representation in human infants. *J Acoust Soc Am* 137(6):3346–3355.
- Behroozmand R, Oya H, Nourski KV, Kawasaki H, Larson CR, Brugge JF, Howard MA 3rd, Greenlee JD. (2016) Neural correlates of vocal production and motor control in human Heschl's gyrus. *J Neurosci* 36(7):2302–2315.
- Courant R, Robbins H. (1996) *What Is Mathematics? An Elementary Approach to Ideas and Methods*. 2nd ed. New York, NY: Oxford University Press.
- Dillon H. (2001) *Hearing Aids*. New York, NY: Thieme.
- Feigin JA, Kopun JG, Stelmachowicz PG, Gorga MP. (1989) Probe-tube microphone measures of ear-canal sound pressure levels in infants and children. *Ear Hear* 10:254–258.
- Galbraith GC, Amaya EM, de Rivera JM, Donan NM, Duong MT, Hsu JN, Tran K, Tsang LP. (2004) Brain stem evoked response to forward and reversed speech in humans. *Neuroreport* 15(13):2057–2060.
- Galbraith GC, Bagasan B, Sulahian J. (2001) Brainstem frequency-following response recorded from one vertical and three horizontal electrode derivations. *Percept Mot Skills* 92(1):99–106.
- Goldstein LJ, Schneider DI, Lay DC, Asmar NH. (2009) *Calculus and Its Applications*. 12th ed. Upper Saddle River, NJ: Prentice-Hall.
- Hornickel J, Skoe E, Zecker S, Kraus N. (2009) Subcortical differentiation of stop consonants relates to reading and speech-in-noise perception. *Proc Natl Acad Sci USA* 106:13022–13027.
- Jeng F-C, Chung H-K, Lin C-D, Dickman B, Hu J. (2011) Exponential modeling of human frequency-following responses to voice pitch. *Int J Audiol* 50(9):582–593.
- Jeng F-C, Hu J, Dickman B, Montgomery-Reagan K, Tong M, Wu G, Lin CD. (2011) Cross-linguistic comparison of frequency-following responses to voice pitch in American and Chinese neonates and adults. *Ear Hear* 32(6):699–707.
- Jeng F-C, Peris KS, 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(2):456–465.
- Jeng F-C, Schnabel EA, Dickman BM, Hu J, Li X, Lin CD, Chung HK. (2010) Early maturation of frequency-following responses to voice pitch in infants with normal hearing. *Percept Mot Skills* 111(3):765–784.
- Joint Committee on Infant Hearing. (1994) Joint committee on infant hearing 1994 position statement. *ASHA* 36(12):38–41.
- Keefe DH, Bulen JC, Hoberg Arehart K, Burns EM. (1993) Ear-canal impedance and reflection coefficient in human infants and adults. *J Acoust Soc Am* 94:2617–2638.
- Krishnan A, Xu Y, Gandour JT, Cariani PA. (2004) Human frequency-following response: representation of pitch contours in Chinese tones. *Hear Res* 189(1–2):1–12.
- Krishnan A, Xu Y, Gandour J, Cariani P. (2005) Encoding of pitch in the human brainstem is sensitive to language experience. *Brain Res Cogn Brain Res* 25(1):161–168.
- Miller CA, Abbas PJ, Robinson BK, Nourski KV, Zhang F, Jeng F-C. (2006) Electrical excitation of the acoustically sensitive auditory nerve: single-fiber responses to electric pulse trains. *J Assoc Res Otolaryngol* 7(3):195–210.
- 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 USA* 104:15894–15898.
- Nourski KV, Abbas PJ, Miller CA, Robinson BK, Jeng F-C. (2005) Effects of acoustic noise on the auditory nerve compound action potentials evoked by electric pulse trains. *Hear Res* 202(1–2):141–153.
- Nourski KV, Brugge JF, Reale RA, Kovach CK, Oya H, Kawasaki H, Jenison RL, Howard MA. (2013) Coding of repetitive transients by auditory cortex on posterolateral superior temporal gyrus in humans: an intracranial electrophysiology study. *J Neurophysiol* 109(5):1283–1295.
- Rubel EW, Ryals BM. (1983) Development of the place principle: acoustic trauma. *Science* 219(4584):512–514.
- Russo N, Nicol T, Musacchia G, Kraus N. (2004) Brainstem responses to speech syllables. *Clin Neurophysiol* 115(9):2021–2030.
- Russo NM, Nicol TG, Zecker SG, Hayes EA, Kraus N. (2005) Auditory training improves neural timing in the human brainstem. *Behav Brain Res* 156(1):95–103.
- Russo NM, Skoe E, Trommer B, Nicol T, Zecker S, Bradlow A, Kraus N. (2008) Deficient brainstem encoding of pitch in children with autism spectrum disorders. *Clin Neurophysiol* 119(8):1720–1731.
- Scollie SD, Seewald RC, Cornelisse LE, Jenstad LM. (1998) Validity and repeatability of level-independent HL to SPL transforms. *Ear Hear* 19(5):407–413.
- 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. (2015) Stability and plasticity of auditory brainstem function across the lifespan. *Cereb Cortex* 25(6):1415–1426.
- Song JH, Skoe E, Wong PCM, Kraus N. (2008) Plasticity in the adult human auditory brainstem following short-term linguistic training. *J Cogn Neurosci* 20(10):1892–1902.
- Swaminathan J, Krishnan A, Gandour JT. (2008) Pitch encoding in speech and nonspeech contexts in the human auditory brainstem. *Neuroreport* 19:1163–1167.
- Wong PCM, Skoe E, Russo NM, Dees T, Kraus N. (2007) Musical experience shapes human brainstem encoding of linguistic pitch patterns. *Nat Neurosci* 10:420–422.