

Research Article

Development of Phase Locking and Frequency Representation in the Infant Frequency-Following Response

Katlyn B. Van Dyke,^a Rachel Lieberman,^a Alessandro Presacco,^{a,b} and Samira Anderson^{a,b,c}

Purpose: This study investigates the development of phase locking and frequency representation in infants using the frequency-following response to consonant–vowel syllables.

Method: The frequency-following response was recorded in 56 infants and 15 young adults to 2 speech syllables (/ba/ and /ga/), which were presented in randomized order to the right ear. Signal-to-noise ratio and F_{sp} analyses were used to verify that individual responses were present above the noise floor. Thirty-six and 39 infants met these criteria for the /ba/ or /ga/ syllables, respectively, and 31 infants met the criteria for both syllables. Data were analyzed to obtain measures of phase-locking strength and spectral magnitudes.

Results: Phase-locking strength to the fine structure in the consonant–vowel transition was higher in young adults than in infants, but phase locking was equivalent at the fundamental frequency between infants and adults. However, frequency representation of the fundamental frequency was higher in older infants than in either the younger infants or adults.

Conclusion: Although spectral amplitudes changed during the first year of life, no changes were found with respect to phase locking to the stimulus envelope. These findings demonstrate the feasibility of obtaining these measures of phase locking and fundamental pitch strength in infants as young as 2 months of age.

Language development depends on strong neural representation of phonemes. Many studies to date have used electrophysiological methods to examine the relationship between language development and how the brain processes varying components of speech. Most studies examining the link between neural representation of speech and later language development in infants have used cortical auditory-evoked potentials. However, recent work with preschoolers and school-age children has suggested that the *frequency-following response* (FFR) may be a viable tool for predicting later language development (White-Schwoch et al., 2015). The FFR is a noninvasive, objective measure of auditory coding generated primarily in the midbrain that also reflects contributions from the auditory nerve, brainstem, and cortex (Chandrasekaran & Kraus,

2010; Coffey, Herholz, Chepesiuk, Baillet, & Zatorre, 2016; Hoormann, Falkenstein, Hohnsbein, & Blanke, 1992; Smith, Marsh, & Brown, 1975). This response has historically been elicited by pure tones; however, recent studies have measured FFRs to more complex auditory stimuli, such as music and speech (Basu, Krishnan, & Weber-Fox, 2010; Mamo, Grose, & Buss, 2016; Shahin, Trainor, Roberts, Backer, & Miller, 2010; Skoe & Kraus, 2010). Furthermore, the FFR can be used to assess phase-locked activity to the formants of vowel stimuli (Krishnan, 2002; Won et al., 2016).

The FFR reflects the summed activity that phase locks to the frequency spectrum of the stimulus up to 1000 Hz. This phase locking represents the temporal structure of the stimulus (Galbraith, Arbagey, Branski, Comerci, & Rector, 1995). Therefore, it is possible to analyze the integrity of temporal processing of the midbrain using the FFR (Anderson, Parbery-Clark, White-Schwoch, & Kraus, 2013). Spectral information is also well represented in the FFR, with clear peaks at the first and second formants (Krishnan, 2002; Krishnan, Xu, Gandour, & Cariani, 2005).

Knowledge of the development of auditory phase locking may contribute to understanding the time course of speech and language development. Phase locking has been demonstrated to both the temporal fine structure (TFS) and the envelope of the stimulus in the auditory

^aDepartment of Hearing and Speech Sciences, University of Maryland, College Park

^bProgram in Neuroscience and Cognitive Science, University of Maryland, College Park

^cLanguage Science Center, University of Maryland, College Park

Correspondence to Samira Anderson: sander22@umd.edu

Editor: Frederick (Erick) Gallun

Associate Editor: Suzanne Purdy

Received June 18, 2016

Revision received November 23, 2016

Accepted April 8, 2017

https://doi.org/10.1044/2017_JSLHR-H-16-0263

Disclosure: The authors have declared that no competing interests existed at the time of publication.

nerve of the chinchilla (Javel, 1980) and in the FFR in humans (Krishnan, 1999). Animal models show that phase locking develops first in the low frequencies and later in the high frequencies. Although phase locking in the kitten auditory nerve and brainstem emerges rapidly within 7 to 10 postnatal days for low frequencies, adultlike phase locking for high frequencies does not occur until the animals reach 3 weeks of age (Brugge, Javel, & Kitzes, 1978; Kettner, Feng, & Brugge, 1985). The earlier development of low-frequency phase locking has also been demonstrated in midbrain structures (Pierson & Snyder-Keller, 1994; Romand & Ehret, 1990). Changes in synaptic transmission and increases in myelination may account for improved temporal precision with development (Brenowitz & Trussell, 2001; Sanes, 1993; Venkataraman & Bartlett, 2013).

The development of frequency and temporal representation in infants may follow the same patterns demonstrated in animal studies. There is a rapid increase in strength of fundamental frequency (F_0) representation in the first 3 months of life (Jeng et al., 2016), but the neural representation of higher frequency spectral content may have a more protracted course of development (Anderson, Parbery-Clark, White-Schwoch, & Kraus, 2015). Coincident with changes in myelination (Moore, Perazzo, & Braun, 1995; Sano, Kaga, Kuan, Ino, & Mima, 2007), peak latencies to click and speech stimuli decrease dramatically in infancy and continue to decrease until early childhood (Anderson et al., 2015; Skoe, Krizman, Anderson, & Kraus, 2015).

A number of studies have demonstrated the feasibility of recording FFRs in infants. Adultlike amplitudes were first observed in newborn responses to 10-ms tone bursts (Gardi, Salamy, & Mendelson, 1979). More recent studies have investigated infant responses to speech stimuli. Representation of the F_0 has been demonstrated in newborn infant responses to lexical tones (Jeng et al., 2011; Jeng et al., 2016). Furthermore, replicable FFRs to a 40-ms consonant–vowel (CV) /da/ syllable can be elicited in infants as young as 3 months of age (Anderson et al., 2015).

With the aim of investigating changes in phase locking and spectral representation to syllables in infants during the first year of life, this study recorded responses to 120-ms /ba/ and /ga/ stop-consonant syllables in 56 infants. The 40-ms stimulus used in the Anderson et al. (2015) study included the formant transition only, not the steady-state region of the vowel. The current study expands on the Anderson et al. (2015) study in several ways. The Anderson et al. (2015) study recorded responses to a speech token that contained a 30-ms formant transition without a steady-state vowel and confined analyses to the temporal envelope. The current study used longer stimuli containing both 50-ms formant transition and 60-ms steady-state vowel regions, which are long enough to obtain meaningful analyses of the response to the TFS in addition to the envelope. Second, this study expanded the age range to include both younger and older infants (from 2 to 12 months) to allow for a more complete picture of developmental changes. Last, these responses were also obtained in young adults to provide a reference for the time course of development of neural synchrony and

response amplitudes, as represented by phase-locking strength and spectral magnitudes.

Based on the protracted course of development noted for representation of spectral and temporal characteristics of speech in behavioral and electrophysiological studies (Anderson et al., 2015; Benson Spetner & Werner Olsho, 1990; Werner, 1999), we hypothesized that responses of young adults will show higher spectral magnitudes and stronger phase locking for the harmonics corresponding to the first formant of speech syllables than in infants. Furthermore, given early maturation of neural representation to voice pitch (Jeng et al., 2016; Jeng et al., 2010), it is hypothesized that phase locking to the F_0 will be equivalent in infants and young adults.

Method

Participants

Fifty-six infants, ages 2 to 12 months, were recruited from the University of Maryland Infant and Child Studies Consortium database: All of the infants were from families in Washington, DC and Maryland. Of these 56, nine infants could not complete the testing due to excessive crying or movement that contaminated the responses, and four were removed from the analysis because they failed to meet response signal-to-noise ratio (SNR) and F_{sp} criteria. The final group of infants comprised 43 infants (M age = 193.90 ± 82.09 days; 26 boys, 17 girls). Infants were placed into two separate groups for comparison: younger infants from ages 2 to 7 months ($n = 21$, M age = 124.61 ± 35.54 days; 15 boys, six girls), and older infants from ages 7 months, 1 day to 1 year ($n = 22$, M age = 263.19 ± 49.72 days; 11 boys, 11 girls). Data were also collected in 15 young adults (M age = 21.53 ± 1.73 years; five men, 10 women). All infants reportedly passed their newborn hearing screening and had no familial history of hearing loss, normal developmental history, and unremarkable otologic and neurologic histories. Normal middle ear function was verified by performing immittance tympanometry in the infants (Type A tympanogram, 1000-Hz probe tone, IMP440, Interacoustics A/S, Denmark). Normal outer hair cell function was confirmed by measuring distortion-product otoacoustic emissions in the infants (+6 dB SNR at 4/5 frequencies from 2 to 8 kHz, DPOAE440, Interacoustics A/S, Denmark). The young adults had normal audiometric thresholds (≤ 20 dB HL, 0.125 to 8 kHz). Auditory brainstem responses (ABRs) to click stimuli were obtained to verify neural integrity in the right ear in both groups (normal Wave V latencies for gestational age, Intelligent Hearing System SmartEP system, IHS, Miami, FL). Participants were compensated for their time, and the parents of babies also received a baby book of their choice. The procedures were approved by the University of Maryland Institutional Review Board.

Electrophysiology Stimuli

Two speech syllables, /ba/ and /ga/, were synthesized in Praat at a 20-kHz sampling rate using a Klatt-based

synthesizer (Boersma & Weenink, 2009). The duration of both stimuli was 120 ms, with voicing (F_0 : 100 Hz) onset at 10 ms and a 50-ms transition from the consonant to the vowel. The vowel was sustained for 60 ms. After an initial 10-ms stop burst, the first three formant frequencies shifted during the 50-ms transition. For /ba/, the second formant (F_2) onset was at 900 Hz and for /ga/, at 2480 Hz; these formants shifted to 1240 Hz and were equal for both syllables during the vowel region of the response. The remaining formants were identical, with the first formant shifting from 400 to 720 Hz and the third formant shifting from 2580 to 2500 Hz, and formants F_4 , F_5 , and F_6 steady through the duration of the stimuli (F_4 at 3300 Hz, F_5 at 3750 Hz, and F_6 at 4900 Hz).

Recording Procedure

The /ba/ and /ga/ stimuli were presented monaurally to the right ear using the IHS SmartEP Continuous Acquisition Module (Intelligent Hearing Systems, Miami, FL) with alternating polarities at 80 dB SPL at a rate of 6.67 sweeps/s through electromagnetically shielded insert earphones. This rate was chosen to minimize testing time and was based on a previous study that showed relatively short interstimulus intervals (35 to 45 ms) were sufficient for stimuli of 250-ms duration (Jeng & Warrington, 2011). Prior to each recording, we calibrated the stimuli to 80 dB SPL using a Larson Davis 824S Type 1 sound level meter (PCB Piezotronics, Inc., Depew, NY) coupled to an earphone adaptor. We sampled the stimuli over 60 s to obtain the average sound pressure level. The relatively high stimulation rate was chosen to minimize test time with infants. A montage of three electrodes (Cz active, forehead ground, right earlobe reference) was used. Two blocks of 1,500 artifact-free sweeps were collected for both /ba/ and /ga/, presented in randomized order. During the recording session, infants reclined on their mothers' laps in a sleeping or drowsy state while the mothers sat in a recliner.

Data Reduction

Responses were sampled at 20 kHz and bandpass filtered offline from 70 to 2000 Hz with a Butterworth filter (Butterworth, 1930) and epoched from -20 to 140 ms referenced to the stimulus onset. This frequency range was selected to minimize low-frequency oscillations from the cortex while maximizing SNR. SNR in dB was calculated using the following formula:

$$20 * \log_{10}(\text{RMS Post-Stimulus} / \text{RMS Noise})$$

where RMS is the root mean square, the poststimulus period is defined as 5–120 ms, and the prestimulus period (noise) is defined as -20 to 0 ms.

The artifact rejection criterion was set at 30 mV. The final response for each syllable was computed for the first 2,500 artifact-free sweeps. The spectrograms and waveforms corresponding to the stimuli with average response waveforms for the two infant groups and the young adult

group are displayed in Figure 1. The transition and steady-state regions of the response, 20–60 ms and 60–120 ms, respectively, are noted on the waveforms. These regions correspond to transition and steady-state regions of the stimuli but are shifted by 10 ms to account for neural transmission delay.

The alternating polarities were added to maximize the response to the temporal envelope and subtracted to maximize the response to the TFS (Aiken & Picton, 2008; Campbell, Kerlin, Bishop, & Miller, 2012; Krishnan, 2002). Added responses were averaged and used to assess the periodic components of the response through the phase-locking factor (PLF) and magnitude of the F_0 . Subtracted responses were averaged and then used to analyze the PLF and spectral energy of the harmonics contained in the first formant.

F_{sp} Analysis

An F_{sp} analysis was performed to verify the quality of responses included in the analysis. The F_{sp} was calculated according to the following formula:

$$F_{sp} = \text{VAR}(S) / \text{VAR}(SP)$$

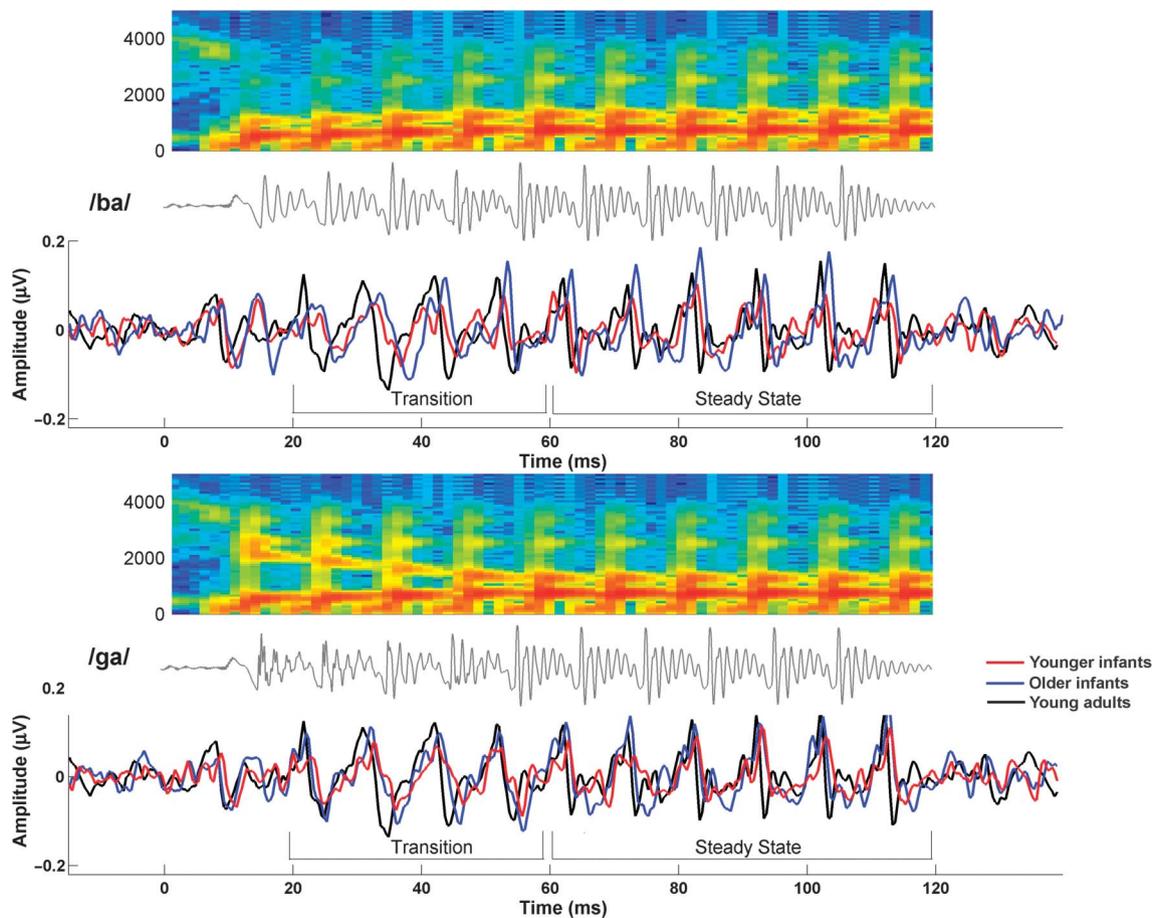
where VAR(S) refers to the variance in the response amplitude corresponding to the response region (20–120 ms), and VAR(SP) refers to the variance of a single point in the response divided by the number of sweeps. This method has been proposed in previous studies as an objective method of verifying response detection (Elberling & Don, 1984; Sininger, 1993). Responses with an F_{sp} value less than 3.14 were removed from the analysis. The remaining babies all had responses with SNR dB values > 0.5.

Data Analysis

PLF

Phase locking was investigated using Morlet's wavelets. This method was chosen to optimize the frequency resolution at all the frequencies tested, and it allows for a visual representation of trial-by-trial phase coherence, referred to as PLF, for a specific frequency range at each individual point in time during a response. A procedure similar to that described by Tallon-Baudry, Bertrand, Delpuech, and Pernier (1996) was used. The wavelet function was convoluted with the response at each time point, providing an amplitude (arbitrary units) and phase (radians) value for each Time \times Frequency bin. The amplitude corresponds to signal energy, and the phase corresponds to the direction of the momentary deflection in the signal. The amplitude was normalized to 1, regardless of its voltage. The normalized complex energy was calculated for each sweep and then averaged across trials, leading to a complex value that represents the phase in the time-frequency domain for each subject tested. The resulting PLF is the average phase distribution of the response (average of all trials), with a factor of 1 corresponding to perfect phase, and 0 corresponding to no phase coherence in every single trial. Overall,

Figure 1. The spectrograms and stimulus waveforms for the /ba/ and /ga/ syllables are displayed with the average response waveforms in younger (red) and older (blue) infants and young adults (black). The periodicity of the stimuli, corresponding to the 10-ms period of the 100-Hz F_0 , is mirrored in the responses. The transition (20–60 ms) and steady-state (60–120 ms) time regions are indicated in the bottom panels.



phase coherence is small (0 to .03 for the TFS and 0 to .05 for the envelope) due to the small size of the response generated in the brainstem/midbrain, consistent with previous studies (Anderson, Parbery-Clark, White-Schwoch, & Kraus, 2012; Ruggles, Bharadwaj, & Shinn-Cunningham, 2012). The mean values for the F_0 of the envelope were averaged in 10-Hz bins (95–105 Hz) across the transition and steady-state regions. The mean values of the TFS were averaged in 10-Hz bins for the frequencies that correspond to the first formant (i.e., 495–505 Hz in the transition region of the response and 695–705 Hz in the steady-state region). These values were chosen because the first formant frequency is approximately 500 Hz at the start of the transition region (20 ms) and stabilizes to 700 Hz by the start of the steady-state region (60 ms). The PLF was similarly calculated for the prestimulus region to ensure that the low values we obtained in the responses were above the noise floor. In addition, the wgn function in MATLAB (MathWorks, Natick, MA) was used to generate 50 white Gaussian noise files, and the PLF values obtained from the three participant groups were compared to the PLF calculated for the noise files.

Fast Fourier Transform

Spectral magnitudes in the steady-state and transition regions were computed using the Welch method with zero padding to allow a fast Fourier transform (FFT) bin resolution of 1 Hz for the F_0 (envelope) and for the fifth through seventh harmonics (500–700 Hz, TFS) using 60-Hz bins. To ensure that the measured amplitudes were actual responses (exceeded the noise floor), spectral energy was calculated for the prestimulus period (–20 to 0 ms), and the spectral energy of the steady-state region was compared to the energy of the prestimulus region.

Statistical Analysis

The infants were divided in two age groups: 2 months to 7 months (younger infants), and 7 months, 1 day to 12 months (older infants), with 21 younger infants and 23 older infants. These age groups were selected because ABR thresholds to low-frequency tone bursts reach maturity between 4 and 6 months of age, with higher frequencies reaching maturity within the following months (Marcoux, 2011). Only responses with positive SNRs greater than 0.5 and an

F_{sp} value greater than 3.41 were included in the analysis. Seventeen younger and 19 older infants met the SNR and F_{sp} criteria for /ba/, 19 younger and 20 older infants met the criteria for /ga/, and 14 younger and 17 older infants met the criteria for both syllables. All of the young adults met the SNR and F_{sp} criteria. Univariate analyses of variance (ANOVAs) were performed to compare group differences in the PLF and in the magnitudes of F_0 representation in the envelope and first formant representation in the TFS (500 Hz, transition; 700 Hz, steady state). A two-way repeated-measures ANOVA (within-subjects: syllable, two levels; region; two levels; between-subjects: group, three levels) was performed to test for within-subjects differences for the syllable and region and between-subjects for differences in phase-locking strength and spectral magnitudes.

Results

PLF Envelope

Transition Region

The PLF of the F_0 was calculated in response to the envelope of the /ba/ and /ga/ syllables. There was no main effect of group for the /ba/, $F(2, 50) = 1.013, p = .371, \eta_p^2 = .041,$

but there was a significant effect of group for the /ga/, $F(2, 53) = 3.502, p = .038, \eta_p^2 = .121$. Post hoc analyses showed that the PLF was greater in the young adults compared to the older infants ($p = .043$) in response to the /ga/, but the differences between the two groups of infants and between the young adults and younger infants were not significant (all p values $> .05$). To ensure that the PLF values exceeded the noise floor, the PLF of the F_0 was also calculated for the prestimulus region (-20 to 0 ms), and these values were compared to those obtained for the transition region. The PLF was significantly higher in the transition region compared to the prestimulus region for both the /ba/ and /ga/ syllables (/ba/: $F[1, 48] = 27.073, p < .001, \eta_p^2 = .361$; /ga/: $F[1, 51] = 63.099, p < .001, \eta_p^2 = .553$). The Group \times Region interactions for both syllables were not significant ($p > .1$), indicating that the region differences were similar across groups.

Steady-State Region

No differences between the three groups were observed for the /ba/ syllable, $F(2, 50) = 0.935, p = .400, \eta_p^2 = .037,$ or the /ga/ syllable, $F(2, 53) = 0.147, p = .864, \eta_p^2 = .006$. The PLF for the envelope is displayed in Figures 2 and 3 for /ba/ and /ga/, respectively.

Figure 2. Left panels: Average phase locking from individual responses to the /ba/ envelope is represented in the time–frequency domain, with darker red colors indicating higher phase-locking frequency (PLF). Right panels: Line graphs depicting the PLF for the F_0 and second harmonic in younger infants (red), older infants (blue), young adults (black), and white Gaussian noise (dashed gray). There is no difference in the PLF between the three groups for the transition and steady-state regions at F_0 (100 Hz) or the second harmonic (200 Hz). Note that the PLF is significantly above the noise floor. Error bars = 1 SE.

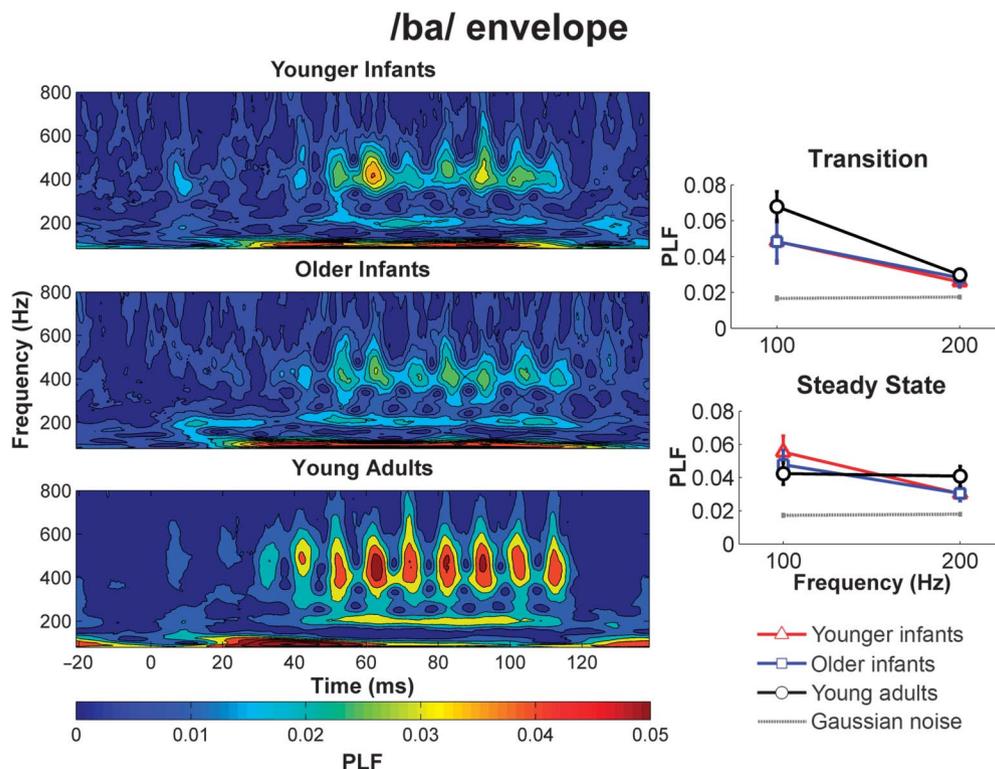
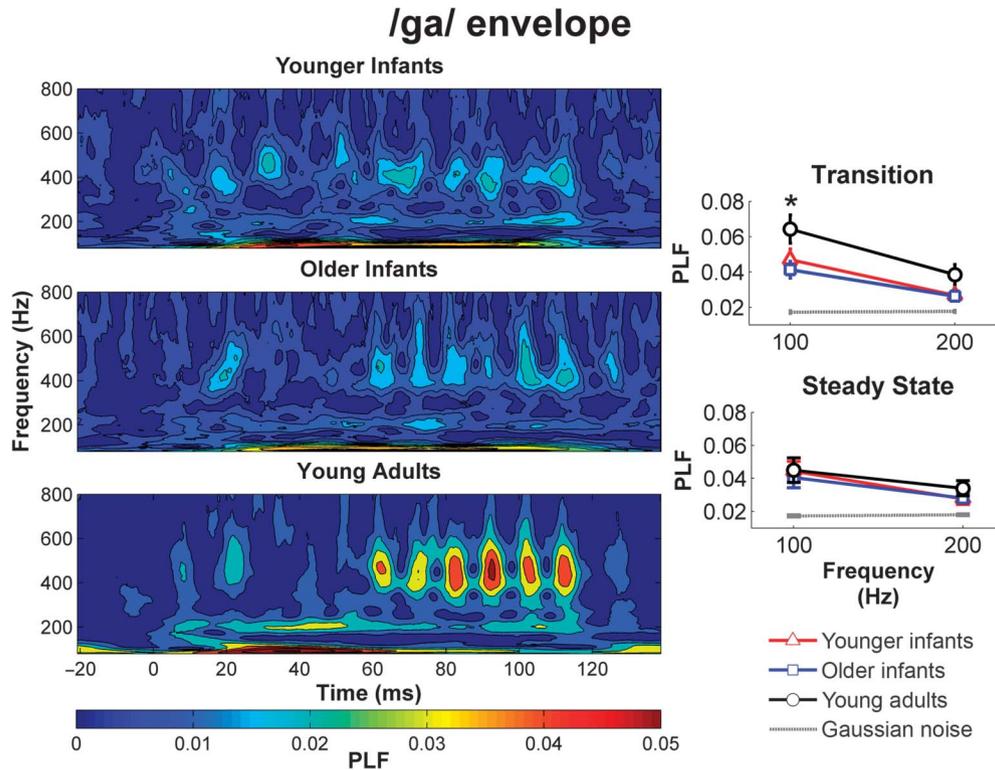


Figure 3. Left panels: Average phase locking from individual responses to the /ga/ envelope is represented in the time–frequency domain, with darker red colors indicating higher phase-locking frequency (PLF). Phase locking is greater in the young adults than in either infant group at the F_0 . Right panels: Line graphs depicting the PLF for the F_0 and second harmonic in younger infants (red), older infants (blue), young adults (black), and white Gaussian noise (dashed gray). Transition: The young adults have a significantly higher PLF value at the F_0 only. Steady state: The PLF is equivalent between groups for the F_0 . * $p < .05$. Note that the PLF is significantly above the noise floor. Error bars = 1 SE.



TFS

Transition Region

The PLF was calculated for 500 Hz (transition first formant) in response to the TFS of the /ba/ and /ga/ syllables. There was a main effect of group, $F(2, 50) = 7.802$, $p = .001$, $\eta_p^2 = .245$, for the /ba/ transition. Post hoc t tests showed that this effect was driven by group differences between the young adults and both groups of infants (younger: $p = .002$; older: $p = .014$). No differences were found between the younger and older infants ($p = .701$). In response to the /ga/ transition there was a main effect of group, $F(2, 53) = 10.7007$, $p < .001$, $\eta_p^2 = .296$, with post hoc t tests showing that the PLF is higher in young adults compared to both groups of infants (younger: $p = .002$; older: $p < .001$) and no differences between younger and older infants ($p = .824$). There was a main effect of syllable, $F(1, 46) = 2.459$, $p = .049$, with higher PLF values for the /ba/ syllable than the /ga/ syllable. Last, the PLF was significantly higher in the transition region compared to the prestimulus region for both the /ba/ and /ga/ syllables (/ba/: $F[1, 48] = 61.856$, $p < .001$, $\eta_p^2 = .563$; /ga/: $F[1, 51] = 68.781$, $p < .001$, $\eta_p^2 = .574$). The Group \times Region interactions for both syllables were not significant ($p > .1$),

indicating that the region differences were similar across groups.

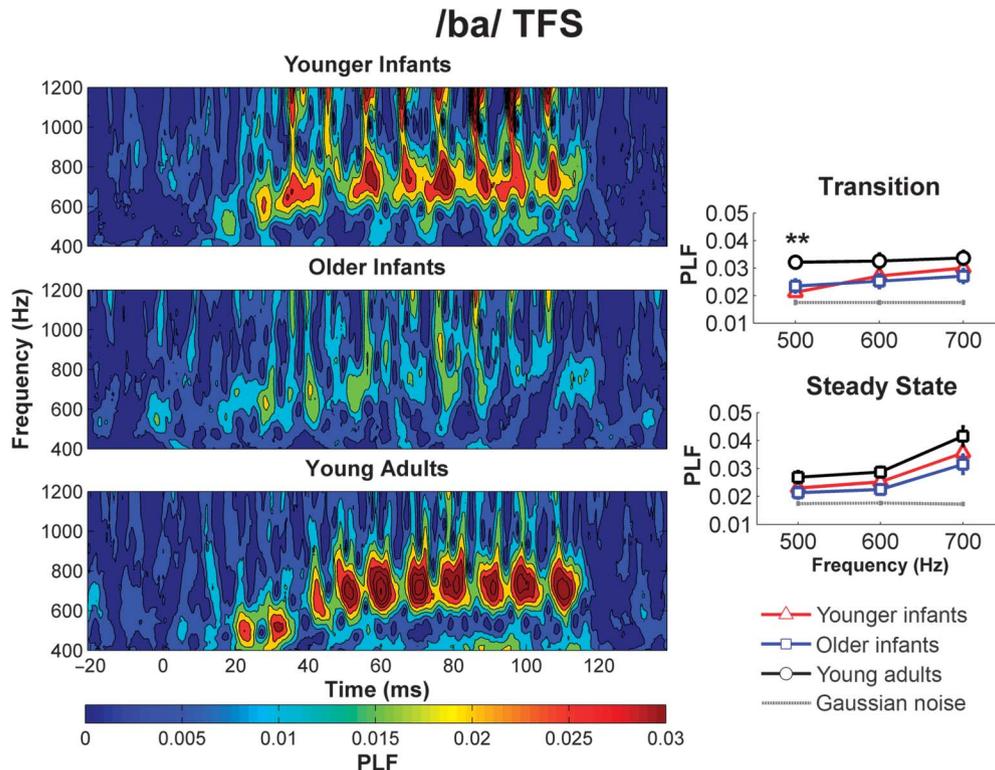
Steady-State Region

The PLF was calculated for 700 Hz (steady-state first formant). No effect of group was observed for the /ba/ steady state, $F(2, 50) = 1.787$, $p = .178$, $\eta_p^2 = .069$, or the /ga/ steady state, $F(2, 53) = 1.029$, $p = .365$, $\eta_p^2 = .039$. The PLF for the TFS is displayed in Figures 4 and 5 for /ba/ and /ga/, respectively.

Syllable and Region Comparison

Responses were compared within individuals and across groups to determine if the FFR reflects the acoustic differences in the syllables. Because the F_0 did not differ between the syllables, this analysis was performed only on the TFS. The PLF was higher for the /ba/ transition than for the /ga/ transition, $F(2, 46) = 8.992$, $p = .004$, $\eta_p^2 = .170$, across groups, and there was no Group \times Syllable interaction, $F(2, 46) = 0.974$, $p = .386$, $\eta_p^2 = .042$, suggesting similar effects of syllables across groups. In the steady-state region where the formants are identical, there were no

Figure 4. Left panels: Average phase locking from individual responses to the /ba/ temporal fine structure (TFS) is represented in the time–frequency domain, with darker red colors indicating higher phase-locking frequency (PLF). Phase locking is greater in the young adults than either infant group in the transition to the first formant (500 Hz, H₅). Right panels: Line graphs depicting the PLF for the higher harmonics (H₅–H₇) in younger infants (red), older infants (blue), young adults (black), and white Gaussian noise (dashed gray). Transition: The young adults have significantly higher PLF values than the infants at 500 Hz. Steady state: The PLF is equivalent between groups for 700 Hz (H₇). Note that the PLF is significantly above the noise floor. ***p* < .01. Error bars = 1 SE.



significant differences between syllables, $F(2, 46) = 0.548$, $p = .463$, $\eta_p^2 = .012$. There was no Group \times Syllable \times Region interaction, $F(2, 46) = 1.006$, $p = .374$, $\eta_p^2 = .044$. Mean PLF values and standard deviations for the F₀ and spectral magnitudes for the transition and steady-state regions can be found in Table 1.

FFT Envelope

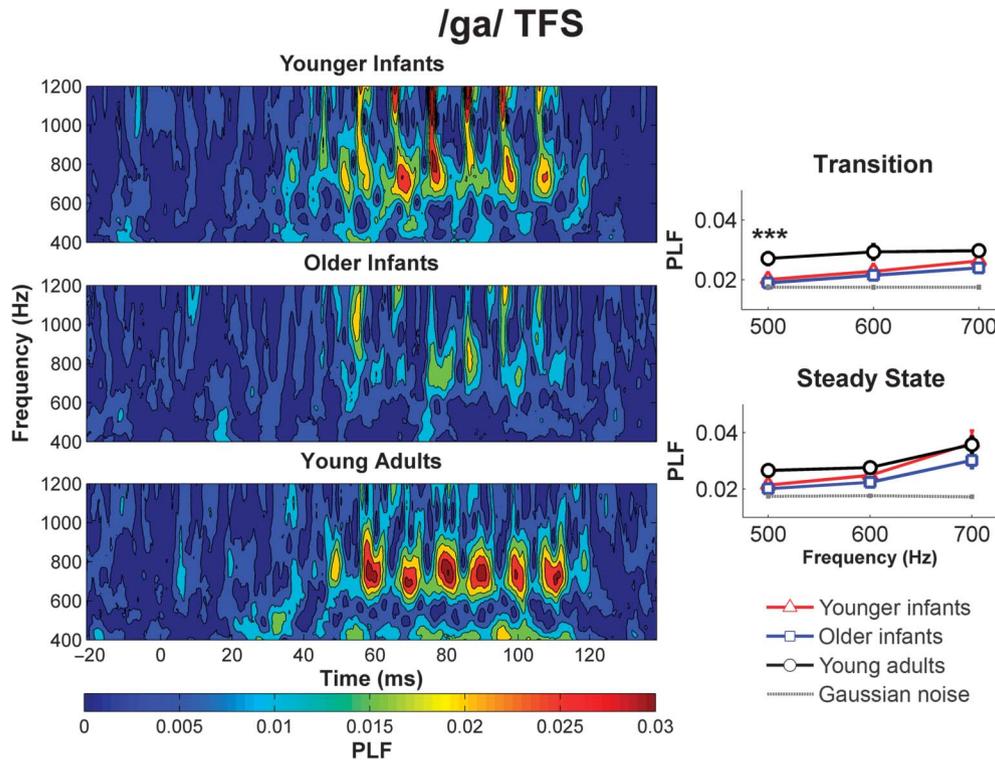
The F₀ magnitude was calculated for the transition and steady state. In the transition, a small main effect of group was found for the /ba/, $F(2, 50) = 3.392$, $p = .042$, $\eta_p^2 = .124$, but post hoc *t* tests were not significant for any group comparison (all *p* values > .05). No main effect of group was found for the /ga/ transition, $F(2, 53) = 2.419$, $p = .099$, $\eta_p^2 = .087$. In the steady-state region of the /ba/ syllable, there was a main effect of group, $F(2, 50) = 7.197$, $p = .002$, $\eta_p^2 = .231$, with post hoc *t* tests showing higher representation of the F₀ in older infants than in the young adults ($p = .002$), but the differences were not significant between the two groups of infants or between the young adults and younger infants (all *p* values > .05). In a similar manner, a main effect of group was found for the /ga/, $F(2, 53) = 5.585$, $p = .006$, $\eta_p^2 = .180$, with post hoc *t* tests

showing higher representation of the F₀ in older infants than in the young adults ($p = .011$), but the differences were not significant between the two groups of infants or between the young adults and younger infants (all *p* values > .05; see Figure 6A and 6B).

TFS

Spectral magnitudes were calculated for 500 Hz for the transition and for 700 Hz for the steady state. A main effect of group was found for the /ba/ transition, $F(2, 50) = 26.370$, $p < .001$, $\eta_p^2 = .524$. This effect was driven by higher spectral magnitudes in the young adults compared to both infant groups ($p < .001$ for both groups), but no significant differences were found between infant groups ($p = .996$). A main effect of group was also found for the /ga/ transition, $F(2, 53) = 4.585$, $p = .015$, $\eta_p^2 = .152$. In this case, post hoc testing revealed that the older infants had higher spectral magnitudes than the younger infants ($p = .042$) and the young adults ($p = .045$), but the younger infants and young adults did not differ ($p = .992$). In the steady state, there was no main effect of group for the /ba/ syllable, $F(2, 50) = 0.432$, $p = .652$, $\eta_p^2 = .018$. There was a small main effect for the /ga/ syllable, $F(2, 53) = 3.461$,

Figure 5. Left panels: Average phase locking from individual responses to the /ga/ temporal fine structure (TFS) is represented in the time–frequency domain, with darker red colors indicating higher phase-locking frequency (PLF). Phase locking is greater in the young adults than in either infant group in the transition region. Right panels: Line graphs depicting the PLF for the higher harmonics (H₅–H₇) in younger infants (red), older infants (blue), young adults (black), and white Gaussian noise (dashed gray). Transition: The young adults have significantly higher PLF values at the first formant, 500 Hz. Steady state: The PLF is equivalent between groups at all frequencies. Note that the PLF is significantly above the noise floor. ****p* < .01. Error bars = 1 SE.



$p = .039$, $\eta_p^2 = .119$, but the post hoc *t* tests all had *p* values > .05 (see Figure 6C and 6D).

Syllable and Region Comparison

Responses were compared within individuals and across groups to determine if spectral magnitudes of the FFR reflect the acoustic differences in the syllables. There was no main effect of syllable, $F(2, 46) = 0.014$, $p = .905$, $\eta_p^2 < .001$.

Discussion

The PLF was stronger in the young adults compared to both infant groups, especially for the spectral components of the stimuli; however, phase locking to the F₀ was equivalent in all three age groups. This adultlike representation of the F₀ in infants has been demonstrated previously (Jeng et al., 2016). FFT analyses demonstrated greater spectral magnitudes in the regions corresponding to the first formant of speech (500-Hz transition, 700-Hz steady state) in the young adults compared to both groups of infants, whereas older infants had larger representation of the F₀ compared to the younger infants and young

adults. When comparing differences between stimuli, there was higher phase locking for the /ba/ syllable in the TFS of the transition region; however, no difference was found in the steady-state region of the syllable where the syllables have identical vowels.

Phase Locking

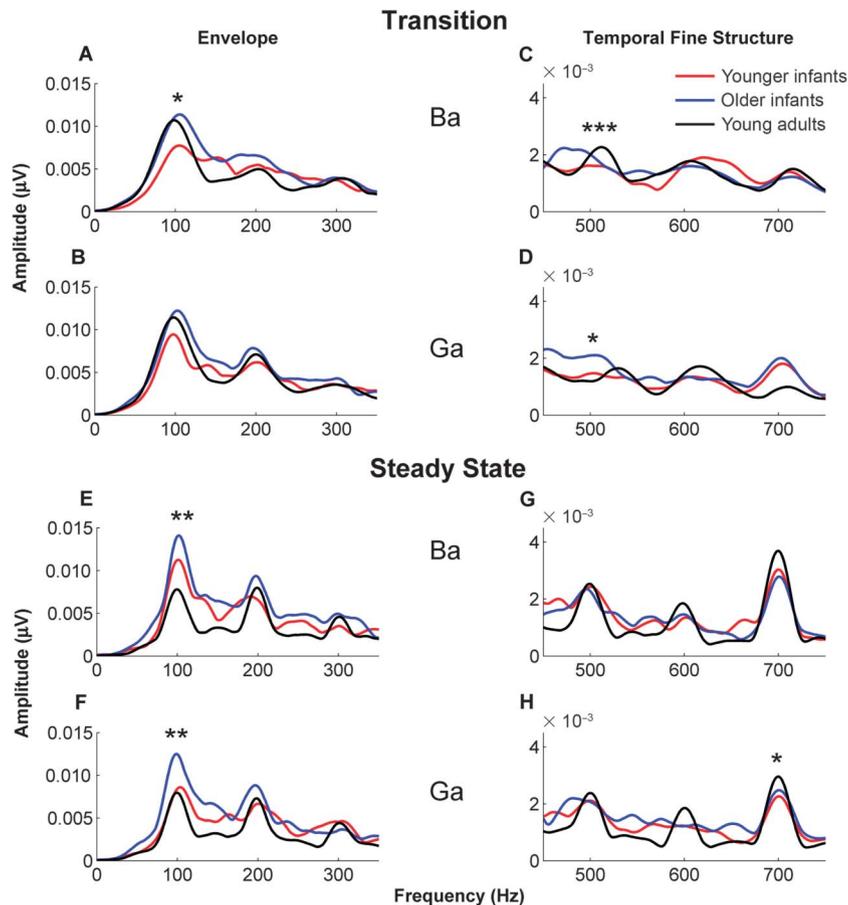
Phase-locking strength in response to the envelope of the /ga/ and /ba/ syllable vowel regions and the /ba/ transition regions did not differ between infants and young adults. Young adults were found to have stronger phase locking for the /ga/ transition region of the envelope, but the effect size was quite small (.12). In contrast, there were significant group differences for both syllables in response to the first formant during the formant transition with effect sizes > .2. These group differences were not noted in the vowel region. The group differences during the transition region may be associated with the development of synaptic efficiency during infancy. Venkataraman and Bartlett (2013) investigated GABAergic connections from the inferior colliculus to the thalamus in rat brain slices from three age groups (prehearing, immediate posthearing, and juvenile) and found that GABA inhibitory postsynaptic potentials

Table 1. $M \pm SD$ PLF values for the F_0 and spectral magnitudes for the transition and steady-state regions of the envelope of the /ba/ and /ga/ syllables for each participant group.

Frequency (Hz)	/ba/ transition			/ba/ steady state			/ga/ transition			/ga/ steady state			Noise transition	Noise steady state
	YI	OI	YA	YI	OI	YA	YI	OI	YA	YI	OI	YA		
Envelope														
F_0 , 100	0.048 ± 0.044	0.048 ± 0.052	0.068 ± 0.032	0.055 ± 0.040	0.048 ± 0.038	0.042 ± 0.025	0.047 ± 0.026	0.041 ± 0.021	0.064 ± 0.031	0.044 ± 0.026	0.040 ± 0.027	0.045 ± 0.028	0.017 ± 0.003	0.017 ± 0.002
200	0.026 ± 0.012	0.028 ± 0.021	0.030 ± 0.012	0.031 ± 0.014	0.030 ± 0.020	0.041 ± 0.023	0.027 ± 0.012	0.026 ± 0.010	0.039 ± 0.022	0.028 ± 0.014	0.028 ± 0.010	0.034 ± 0.017	0.017 ± 0.003	0.017 ± 0.002
Temporal fine structure														
500	0.021 ± 0.006	0.023 ± 0.09	0.032 ± 0.007	0.023 ± 0.006	0.021 ± 0.008	0.027 ± 0.008	0.020 ± 0.005	0.019 ± 0.002	0.027 ± 0.006	0.021 ± 0.006	0.020 ± 0.002	0.027 ± 0.008	0.017 ± 0.003	0.017 ± 0.002
600	0.027 ± 0.012	0.025 ± 0.010	0.033 ± 0.010	0.025 ± 0.008	0.022 ± 0.006	0.029 ± 0.007	0.023 ± 0.006	0.021 ± 0.004	0.029 ± 0.008	0.025 ± 0.007	0.022 ± 0.003	0.028 ± 0.009	0.017 ± 0.002	0.017 ± 0.001
700	0.030 ± 0.013	0.027 ± 0.10	0.034 ± 0.009	0.036 ± 0.017	0.031 ± 0.014	0.042 ± 0.013	0.026 ± 0.009	0.024 ± 0.005	0.030 ± 0.008	0.036 ± 0.019	0.030 ± 0.011	0.036 ± 0.007	0.017 ± 0.002	0.017 ± 0.001

Note. Noise $M \pm SD$ is displayed to show the noise floor. PLF = phase-locking frequency; YI = younger infant; OI = older infant; YA = young adult.

Figure 6. Panels A–B: Average magnitudes in individual responses of the transition region of the responses to /ba/ and /ga/ envelope in younger infants (red), older infants (blue), and young adults (black). The older infants have higher magnitudes for the F_0 (100 Hz) in the envelope of the /ba/ but not the /ga/ response. Panels C–D: Spectral magnitudes of responses to the /ba/ and /ga/ temporal fine structure (TFS) transition. Young adults had higher spectral magnitudes at 500 Hz (first formant) for the /ba/ syllable, but older infants had higher spectral magnitudes for the /ga/ syllable. Panels E–F: Average magnitudes in individual responses of the steady-state region of the responses to /ba/ and /ga/ envelope. Older infants had higher representation of the envelope than the young adults. Panels G–H: Spectral magnitudes of responses to the /ba/ and /ga/ TFS steady state. Young adults had higher spectral magnitudes than the infants. * $p < .05$, ** $p < .01$.



were faster and there was less depression to trains of stimuli. They surmised that these changes suggest increased ability to control excitability with development. A balance of excitatory and inhibitory transmission would be important for accurately encoding the features of the rapidly changing formant transition. Developmental changes in high-frequency phase locking have also been observed in auditory nerve fibers. Phase locking is only observed at frequencies below 600 Hz in young kittens, but the frequency limits of phase locking reach 4000 Hz in adult cats (Kettner et al., 1985). Delayed development of high-frequency compared to low-frequency responses has also been found in the mouse inferior colliculus (Romand & Ehret, 1990).

When comparing responses to syllables, stronger phase locking to the transition region was found in responses to the TFS of the lower frequency /ba/ syllable than of the /ga/. Not surprisingly, there were no differences in the steady-state region of the syllables where the

stimuli were identical. Overall, these results support the idea that the FFR reflects acoustic differences in stimuli. The increase in phase locking from the /ga/ to the /ba/ was similar across groups. Although we might have expected the syllable differences to be exaggerated in the infant groups compared to the young adults, the lack of Group \times Syllable interaction may reflect the fact that we assessed phase locking to the first formant, when in fact the two syllables only differ in the second formant frequency. We did not assess phase locking to the second formant as phase locking in the individual infant responses decreased dramatically at frequencies higher than the first formant (700 Hz). A question remains as to why we found syllable differences for the first formant at all, when no differences in the stimuli exist in that frequency range. We speculate that because this is a response representing populations of neurons, we might be observing neural activity arising from frequencies higher than what were actually analyzed.

FFTs

Anderson et al.'s (2015) study compared the FFR to a 40-ms consonant-transition speech token in younger (3–5 month) and older (6–9 month) infants and found that in the responses to the envelope, the spectral magnitudes in the higher harmonics were greater in the older than the younger infants, but the representation of the F_0 was equivalent between groups. The use of speech tokens containing a steady-state vowel region allowed for examination of developmental changes in spectral magnitudes of the TFS of the vowel region in the current study. The results in response to the envelope were similar to those of Anderson et al. (2015) in that no infant age group differences were found for the F_0 . It is interesting to note that we did find that the older infants had higher PLF values for the F_0 compared to the young adults. The reason for this finding is not clear. However, this finding, in combination with higher spectral magnitudes in young adults, leads us to speculate that infants overshoot and lock onto the F_0 more strongly because the representation of the low frequencies are more mature than representation of high frequencies. Then, as high-frequency encoding matures, there is a change in the balance of phase locking in the midbrain.

Clinical Implications

These results demonstrate that speech-sound encoding can be assessed in infants as young as 2 months of age. Previous studies have demonstrated phase-locked activity in the FFR to the F_0 and the formants of speech in adults (e.g., Galbraith, Bhuta, Choate, Kitahara, & Mullen, 1998; Krishnan, 2002; Won et al., 2016). In this study, the presence of phase locking at the F_0 of the envelope and the first formant of the fine structure in normally developing infants provides a basis against which to compare responses in infants who are at risk for language disorders. In a clinical setting, data collection is more practical in infants who are younger and more likely to remain quiet or asleep through the testing compared to older infants above 6 months of age who may be active or noisy. The data were collected using a clinician-friendly system; therefore, it may be possible to implement screening in young infants who are at risk for developing a language impairment.

Future Directions

Although this study informs us about developmental changes, further work needs to be done to determine the test–retest reliability of these measures before they could be implemented in a clinical setting. We note that the effect sizes were small, even for the strongest group differences for phase locking to the transition first formant. Even though the PLF values in this study were similar to those found in other studies (Anderson et al., 2012), they were indeed low values. The fact that these values were so small may decrease their detectability, reducing their efficacy in a clinical setting. Future studies may consider the use of

other FFR measures to improve sensitivity. For example, a recent study found that responses to a speech syllable in noise in preschool predicts phonological development 1 year later (White-Schwoch et al., 2015), so future studies may consider the inclusion of a noise condition.

Acknowledgments

This work was supported by the University of Maryland–College Park Department of Hearing and Speech Sciences and the National Institutes of Health Grant T32DC-00046 (awarded to Alessandro Presacco). The authors wish to thank the infants and parents who graciously participated in our testing. The authors also thank Maeve Salanger, Rachael Stein, Meg Graves, Eve Kronzek, Jen Chisholm, Courtney Wallace, and Tarah Chibber for their assistance with data collection and Evan Smith and Arielle Abrams for their assistance with data analysis.

References

- Aiken, S. J., & Picton, T. W. (2008). Envelope and spectral frequency-following responses to vowel sounds. *Hearing Research, 245*, 35–47.
- Anderson, S., Parbery-Clark, A., White-Schwoch, T., & Kraus, N. (2012). Aging affects neural precision of speech encoding. *The Journal of Neuroscience, 32*, 14156–14164.
- Anderson, S., Parbery-Clark, A., White-Schwoch, T., & Kraus, N. (2013). Auditory brainstem response to complex sounds predicts self-reported speech-in-noise performance. *Journal of Speech, Language, and Hearing Research, 56*, 31–43.
- Anderson, S., Parbery-Clark, A., White-Schwoch, T., & Kraus, N. (2015). Development of subcortical speech representation in human infants. *The Journal of the Acoustical Society of America, 137*, 3346–3355.
- Basu, M., Krishnan, A., & Weber-Fox, C. (2010). Brainstem correlates of temporal auditory processing in children with specific language impairment. *Developmental Science, 13*, 77–91.
- Benson Spetner, N., & Werner Olsho, L. (1990). Auditory frequency resolution in human infancy. *Child Development, 61*, 632–652.
- Boersma, P., & Weenink, D. (2009). *Praat: Doing phonetics by computer* (Version 5.1.05) [Computer program]. Retrieved from <http://www.praat.org>
- Brenowitz, S., & Trussell, L. O. (2001). Maturation of synaptic transmission at end-bulb synapses of the cochlear nucleus. *The Journal of Neuroscience, 21*, 9487–9498.
- Brugge, J. F., Javel, E., & Kitzes, L. M. (1978). Signs of functional maturation of peripheral auditory system in discharge patterns of neurons in anteroventral cochlear nucleus of kitten. *Journal of Neurophysiology, 41*, 1557–1559.
- Butterworth, S. (1930). On the theory of filter amplifiers. *Experimental Wireless and the Wireless Engineer, 7*, 536–541.
- Campbell, T., Kerlin, J., Bishop, C., & Miller, L. (2012). Methods to eliminate stimulus transduction artifact from insert earphones during electroencephalography. *Ear and Hearing, 33*, 144–150.
- Chandrasekaran, B., & Kraus, N. (2010). The scalp-recorded brainstem response to speech: Neural origins and plasticity. *Psychophysiology, 47*, 236–246.
- Coffey, E. B., Herholz, S. C., Chepesiuk, A. M., Baillet, S., & Zatorre, R. J. (2016). Cortical contributions to the auditory frequency-following response revealed by MEG. *Nature Communications, 7*, 11070.

- Elberling, C., & Don, M. (1984). Quality estimation of averaged auditory brainstem responses. *Scandinavian Audiology*, 13, 187–197.
- Galbraith, G. C., Arbagey, P. W., Branski, R., Comerci, N., & Rector, P. M. (1995). Intelligible speech encoded in the human brain stem frequency-following response. *Neuroreport*, 6, 2363–2367.
- Galbraith, G. C., Bhuta, S. M., Choate, A. K., Kitahara, J. M., & Mullen, T. A., Jr. (1998). Brain stem frequency-following response to dichotic vowels during attention. *Neuroreport*, 9, 1889–1893.
- Gardi, J., Salamy, A., & Mendelson, T. (1979). Scalp-recorded frequency-following responses in neonates. *Audiology*, 18, 494–506.
- Hoormann, J., Falkenstein, M., Hohsbein, J., & Blanke, L. (1992). The human frequency-following response (FFR): Normal variability and relation to the click-evoked brainstem response. *Hearing Research*, 59, 179–188.
- Javel, E. (1980). Coding of AM tones in the chinchilla auditory nerve: Implications for the pitch of complex tones. *The Journal of the Acoustical Society of America*, 68, 133–146.
- Jeng, F.-C., Hu, J., Dickman, B., Montgomery-Reagan, K., Tong, M., Wu, G., & Lin, C.-D. (2011). Cross-linguistic comparison of frequency-following responses to voice pitch in American and Chinese neonates and adults. *Ear and Hearing*, 32, 699–707.
- Jeng, F.-C., Lin, C.-D., Chou, M.-S., Hollister, G. R., Sabol, J. T., Mayhugh, G. N., . . . Wang, C. Y. (2016). Development of subcortical pitch representation in three-month-old Chinese infants. *Perceptual and Motor Skills*, 122, 123–135.
- Jeng, F.-C., Schnabel, E. A., Dickman, B. M., Hu, J., Li, X., Lin, C.-D., & Chung, H.-K. (2010). Early maturation of frequency-following responses to voice pitch in infants with normal hearing. *Perceptual and Motor Skills*, 111, 765–784.
- Jeng, F.-C., & Warrington, R. P. (2011). Effects of silent interval on human frequency-following responses to voice pitch. *The Journal of the Acoustical Society of America*, 130, 2454.
- Kettner, R. E., Feng, J. Z., & Brugge, J. F. (1985). Postnatal development of the phase-locked response to low frequency tones of auditory nerve fibers in the cat. *The Journal of Neuroscience*, 5, 275–283.
- Krishnan, A. (1999). Human frequency-following responses to two-tone approximations of steady-state vowels. *Audiology & Neuro-otology*, 4, 95–103.
- Krishnan, A. (2002). Human frequency-following responses: Representation of steady-state synthetic vowels. *Hearing Research*, 166, 192–201.
- Krishnan, A., Xu, Y., Gandour, J., & Cariani, P. (2005). Encoding of pitch in the human brainstem is sensitive to language experience. *Cognitive Brain Research*, 25, 161–168.
- Marcoux, A. M. (2011). Maturation of auditory function related to hearing threshold estimations using the auditory brainstem response during infancy. *International Journal of Pediatric Otorhinolaryngology*, 75(2), 163–170.
- Mamo, S. K., Grose, J. H., & Buss, E. (2016). Speech-evoked ABR: Effects of age and simulated neural temporal jitter. *Hearing Research*, 333, 201–209.
- Moore, J. K., Perazzo, L. M., & Braun, A. (1995). Time course of axonal myelination in the human brainstem auditory pathway. *Hearing Research*, 87, 21–31.
- Pierson, M., & Snyder-Keller, A. (1994). Development of frequency-selective domains in inferior colliculus of normal and neonatally noise-exposed rats. *Brain Research*, 636, 55–67.
- Romand, R., & Ehret, G. (1990). Development of tonotopy in the inferior colliculus. I. Electrophysiological mapping in house mice. *Developmental Brain Research*, 54, 221–234.
- Ruggles, D., Bharadwaj, H., & Shinn-Cunningham, B. G. (2012). Why middle-aged listeners have trouble hearing in everyday settings. *Current Biology*, 22, 1417–1422.
- Sanes, D. H. (1993). The development of synaptic function and integration in the central auditory system. *Journal of Neuroscience*, 13, 2627–2637.
- Sano, M., Kaga, K., Kuan, C.-C., Ino, K., & Mima, K. (2007). Early myelination patterns in the brainstem auditory nuclei and pathway: MRI evaluation study. *International Journal of Pediatric Otorhinolaryngology*, 71, 1105–1115.
- Shahin, A. J., Trainor, L. J., Roberts, L. E., Backer, K. C., & Miller, L. M. (2010). Development of auditory phase-locked activity for music sounds. *Journal of Neurophysiology*, 103, 218–229.
- Sininger, Y. S. (1993). Auditory brain stem response for objective measures of hearing. *Ear and Hearing*, 14, 23–30.
- Skoe, E., & Kraus, N. (2010). Auditory brain stem response to complex sounds: A tutorial. *Ear and Hearing*, 31, 302–324.
- Skoe, E., Krizman, J., Anderson, S., & Kraus, N. (2015). Stability and plasticity of auditory brainstem function across the lifespan. *Cerebral Cortex*, 25, 1415–1426.
- Smith, J. C., Marsh, J. T., & Brown, W. S. (1975). Far-field recorded frequency-following responses: Evidence for the locus of brainstem sources. *Electroencephalography and Clinical Neurophysiology*, 39, 465–472.
- Tallon-Baudry, C., Bertrand, O., Delpuech, C., & Pernier, J. (1996). Stimulus specificity of phase-locked and non-phase-locked 40 Hz visual responses in human. *The Journal of Neuroscience*, 16, 4240–4249.
- Venkataraman, Y., & Bartlett, E. L. (2013). Postnatal development of synaptic properties of the GABAergic projection from the inferior colliculus to the auditory thalamus. *Journal of Neurophysiology*, 109, 2866–2882.
- Werner, L. A. (1999). Forward masking among infant and adult listeners. *The Journal of the Acoustical Society of America*, 105, 2445–2453.
- White-Schwoch, T., Woodruff Carr, K., Thompson, E. C., Anderson, S., Nicol, T., Bradlow, A. R., . . . Kraus, N. (2015). Auditory processing in noise: A preschool biomarker for literacy. *PLoS Biology*, 13(7), e1002196.
- Won, J. H., Tremblay, K., Clinard, C. G., Wright, R. A., Sagi, E., & Svirsky, M. (2016). The neural encoding of formant frequencies contributing to vowel identification in normal-hearing listeners. *The Journal of the Acoustical Society of America*, 139, 1–11.