Journal Pre-proofs

Auditory neurophysiological development in early childhood: a growth curve modeling approach

Elaine C. Thompson, Ryne Estabrook, Jennifer Krizman, Spencer Smith, Stephanie Huang, Travis White-Schwoch, Trent Nicol, Nina Kraus

Please cite this article as: Thompson, E.C., Estabrook, R., Krizman, J., Smith, S., Huang, S., White-Schwoch, T., Nicol, T., Kraus, N., Auditory neurophysiological development in early childhood: a growth curve modeling approach, *Clinical Neurophysiology* (2021), doi: https://doi.org/10.1016/j.clinph.2021.05.025

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Auditory neurophysiological development in early childhood: a growth curve modeling approach

Elaine C. Thompson[a,b,1], Ryne Estabrook[c], Jennifer Krizman[a,b], Spencer Smith[a,b,2], Stephanie Huang[a], Travis White-Schwoch[a,b], Trent Nicol[a,b], and Nina Kraus[a,b,d,e,f,*]

[a]Auditory Neuroscience Laboratory, Northwestern University, Evanston, IL, USA; [b]Department of Communication Sciences, Northwestern University, Evanston, IL, USA; [c]Department of Psychology, University of Illinois at Chicago, Chicago, IL, USA; [d]Institute for Neuroscience, Northwestern University, Evanston, IL, USA; [e]Department of Neurobiology, Northwestern University, Evanston, IL, USA; [f]Department of Otolaryngology, Northwestern University, Chicago, IL, USA

[1]Now at Emory University School of Medicine, Atlanta, GA, USA
[2]Now at University of Texas at Austin, Austin, TX, USA

Corresponding author:
Nina Kraus
2240 Campus Drive
Evanston, IL 60208, USA
E-mail address: nkraus@northwestern.edu
www.brainvolts.northwestern.edu

Highlights

- Auditory neurophysiological processing of sound matures from ages 3-8 years.
- The frequency following response, a subcortical index of sound processing, becomes faster, more robust, and more consistent with age.
- Growth curve modeling reveals individual differences in auditory maturation are evident overall and over time throughout early childhood.

Abstract

Objective: During early childhood, the development of communication skills, such as language and speech perception, relies in part on auditory system maturation. Because auditory behavioral tests engage cognition, mapping auditory maturation in the absence of cognitive influence remains a challenge. Furthermore, longitudinal investigations that capture auditory maturation within and between individuals in this age group are scarce. The goal of this study is to longitudinally measure auditory system maturation in early childhood using an objective approach.

Methods: We collected frequency-following responses (FFR) to speech in 175 children, ages 3-8 years, annually for up to five years. The FFR is an objective measure of sound encoding that predominantly reflects auditory midbrain activity. Eliciting FFRs to speech provides rich details of various aspects of sound processing, namely, neural timing, spectral coding, and response stability. We used growth curve modeling to answer three questions: 1) does sound encoding change across childhood? 2) are there individual differences in sound encoding? and 3) are there individual differences in the development of sound encoding?

Results: Subcortical auditory maturation develops linearly from 3-8 years. With age, FFRs became faster, more robust, and more consistent. Individual differences were evident in each aspect of sound processing, while individual differences in rates of change were observed for spectral coding alone.

Conclusions: By using an objective measure and a longitudinal approach, these results suggest subcortical auditory development continues throughout childhood, and that different facets of auditory processing follow distinct developmental trajectories.

Significance: The present findings improve our understanding of auditory system development in typically-developing children, opening the door for future investigations of disordered sound processing in clinical populations.

Abbreviations
Comparative Fit Index (CFI), fast Fourier transform (FFT), frequency following response (FFR), fundamental frequency (F0), Generalized Linear Model (GLM), Intraclass Correlation Coefficient (ICC), Model 1 (M1), Model 2 (M2), Model 3 (M3), root-mean-square (RMS), Root Mean Square Error of Approximation (RMSEA), sound pressure level (SPL), Tucker-Lewis index (TLI).

1. Introduction

During childhood, growth is pervasive, both physically and physiologically. Communication skills, such as language and speech perception, expand rapidly. By age 2, most children can combine words into 2-word phrases and up to age 4, sentence length grows an average of one word per year. Then, from 3 to 8 years of age, a child's lexicon grows exponentially, from ~ 150 words to ~23,000 words (Macias and Twyman, 2011; Templin, 1957). Concurrently, sentences become more complex, with the addition of prepositions, adjectives, and adverbs by age 5 (Macias and Twyman, 2011). The development of these skills relies in part on the development of the auditory system. Anatomically, peripheral structures (e.g., cochlea) reach maturity in infancy (Abdala and Keefe, 2006; Eggermont and Moore, 2012; Lavigne-Rebillard and Pujol, 1987), while central auditory pathways (e.g., brainstem to cortex) follow a more protracted trajectory through late childhood and even adolescence (Moore and Linthicum, 2007). How the auditory system develops functionally, however, is an open question. Quantifying this maturation is important for understanding differences in speech and language mastery, and ultimately, for identifying disorders of language development.

One line of research suggests that functional maturation is tethered to structural maturation of the auditory system. That is, maturation of distinct auditory perceptual skills follows trajectories roughly corresponding to development of the auditory structures presumed to support these skills (Moore, 2002; Sanes and Woolley, 2011). For example, frequency resolution becomes adult-like by the first year of life (Spetner and Olsho, 1990; Werner, 1996), which is also around the time the cochlea is structurally mature (Eggermont et al., 1996; Ponton et al., 1992). In contrast, perception of temporal cues (e.g., temporal integration) does not reach adult-like levels until ages 6-10 years (Hartley et al., 2000; Jensen and Neff, 1993; Litovsky, 1997; Wightman et al., 1989), possibly due to prolonged neuronal development within the central auditory system (Moore and Linthicum, 2007).

An alternative hypothesis as to why there are multiple trajectories of functional maturation is that some auditory behavioral tasks call upon fundamental mechanisms of "hearing" (i.e., cochlear function) that reach maturity earlier on, while other tasks recruit additional cognitive processes that continue to develop through adolescence. Indeed, auditory perceptual tasks range in their auditory-system demands and the extent to which cognitive processes are recruited. For example, many auditory behavioral tests require sustained attention (Moore et al., 2010), raising the possibility that the different tests are not indexing differential development of auditory processes, but rather the extent to which these processes engage attention. Further complicating the picture is the fact that cognitive processes undergo their own development in early childhood, so it is also possible these behavioral measures are in fact measuring maturation of cognitive function. To examine auditory development in childhood without the confound of cognitive factors, an objective approach is needed (Sanes and Woolley, 2011).

The frequency following response (FFR), a neurophysiological measure of electrical events generated within and throughout the auditory pathway (Bidelman, 2015; Chandrasekaran and Kraus, 2010; White-Schwoch et al., 2017), serves as an objective index of auditory processing because it does not depend on attention or volition (Krizman and Kraus, 2019; Skoe and Kraus, 2010). Because it does not actively engage cognition, the FFR can investigate auditory maturation without the confounds of cognitive processing. Moreover, eliciting the FFR to complex sounds, such as speech, provides rich insights into distinct aspects of auditory system function, including the fidelity with which acoustic features are processed and the health of the auditory system irrespective of sound engagement (Krizman et al., 2020). These aspects of auditory function are indexed through measures derived from the FFR, including neural timing, spectral coding, response stability, and nonstimulus activity (Krizman and Kraus, 2019; Skoe and Kraus, 2010). Examining these aspects simultaneously and objectively would enable a more complete investigation of sound processing, as each offers a perspective into the many ways the nervous system engages with sound.

For example, two aspects of auditory function, neural timing and spectral coding, reflect the auditory system's representation of temporal and frequency information, respectively, both of which are relevant for communication (Krizman and Kraus, 2019; Skoe and Kraus, 2010). Fluctuations in the temporal and spectral components of speech inform a listener of what was said, where the sound originated, and who spoke the message (Carré et al., 2017). The shape and speed of these fluctuations can also signal emotion—such as a speaker's tone or intonation—and comprise the building blocks of speech, like phonemes or syllables (Frick, 1985).

In contrast, response stability reflects the auditory system's ability to reliably encode a message. During a recording session, thousands of stimulus trials are presented to a participant to generate an averaged response waveform. By correlating subsets of these trials, response stability captures how much the response does or does not change over time, and is therefore thought to reflect the reliability of the auditory system in encoding stimulus features (Hornickel and Kraus, 2013; Krizman and Kraus, 2019). Stable auditory processing is believed to facilitate the sound-to-meaning connections associated with language learning and learning to read; unstable auditory processing is a hallmark of children with language disorders (Hornickel and Kraus, 2013; Otto-Meyer et al., 2018).

Finally, nonstimulus activity, captured during the period of silence between stimulus presentations, provides an index of baseline neural activity when not evoked (Krizman and Kraus, 2019); that is, it measures nonstimulus activity thought to reflect background neural and non-neural noise. Background activity levels are influenced by experience: they are higher in children from socioeconomically impoverished backgrounds (Skoe et al., 2013) and they are lower in expert athletes, even when controlling for myogenic artifact (Krizman et al., 2020). Higher levels of nonstimulus activity may hinder communication skills, such as the ability to perceive speech in noisy environments (Anderson et al., 2012).

It is currently unknown how these properties of sound encoding develop in childhood. Mapping the maturational course of these various neurophysiological measures would provide insight into the multifaceted development of auditory function, and therefore could inform clinical diagnostics and interventions, and ultimately guide clinical decision-making.

Moreover, most of what is known about auditory system development has been identified using a cross-sectional approach. Age-related differences in characteristic response components (N1 and P1) of cortical auditory evoked potentials are evident in childhood (Wunderlich and Cone-Wesson, 2006) and adolescence (Ponton et al., 2000; Sharma et al., 1997). Subcortical responses, measured using the FFR to speech, also differ between age groups and have been observed in comparisons of infants (Anderson et al., 2015; Jeng et al., 2010), young (3-5 yo) and school-age children (8-12 yo) (Johnson et al., 2008), and across the lifespan (Skoe et al., 2015). However, one assumption of cross-sectional research is that between-individual differences also reflect within-individual changes over time. Tracking within-individual changes requires a longitudinal design, yet few have employed this approach to examine auditory neurophysiological development, especially in childhood.

Here, we map auditory neurophysiological maturation during childhood using an objective index of auditory processing and a longitudinal approach. The following research questions were asked:
1) Does auditory encoding of speech change across childhood development?
2) Are there individual differences in speech-sound encoding?
3) Are there individual differences in the development of speech-sound encoding over time?

To answer these questions, we measured FFRs to speech in a cohort of children (n = 175) beginning at age 3 or 4 years and continuing each year for up to five years until they were 7 or 8 years old, culminating in 463 test points. Analysis of the longitudinal data was performed using growth curve modeling, which

5

allowed us to determine general developmental changes, individual differences, and individual differences in rates of change. We hypothesized that distinct aspects of auditory processing continue to develop between 3 and 8 years of age and that the different aspects themselves, as well as their development, vary across individuals. Specifically, we predicted that that as children develop, neurophysiological responses to speech become faster, more robust, more consistent, and less noisy, with variation on these measures seen across individuals and across age.

2. Methods

2.1 Participants

One hundred and seventy-five (n = 175) children, recruited from the Chicago area, were included in the study. Participants were monolingual English speakers with no history of a neurological disorder. All passed a screening of peripheral auditory health (normal otoscopy, Type A tympanograms, and distortion product otoacoustic emissions $\geq$ 6 dB sound pressure level (SPL) above noise floor from 0.5 to 4 kHz) and demonstrated normal click-evoked auditory brainstem responses (wave V latency < 6.00 ms in response to a click presented at 80 dB SPL at 31 Hz). Children provided verbal assent and parents and/or legal guardians provided written consent. All experimental procedures were approved by and carried out in accordance to recognized standards set by the Northwestern University Institutional Review Board. Participants were monetarily compensated for their time.

2.2 Testing

Two age cohorts were tested at an initial visit: 3-year-olds (n = 82) and 4-year-olds (n = 93); following this initial test visit, children returned to the lab every 12 months for up to 5 years (Table 1). To achieve greater specificity in the growth curve model, data were analyzed longitudinally with respect to age at test rather than year of study. Thus, although all analyses considered age as a continuous variable, for illustrative purposes, FFR data have been split into age groups in the figures.

[Table 1 here]

2.3 Electrophysiology

FFRs evoked by the speech syllable [da], (Skoe and Kraus, 2010), were collected using a BioSEMI Active2 recording system and auditory brainstem response module. Recordings were performed in an electrically-shielded and sound-attenuated booth (IAC Acoustics, Bronx, NY, USA), and lasted ~30 minutes. The [da], presented in isolation, was played at alternating polarities to the right ear at 80 dB SPL through electromagnetically-shielded insert earphones (ER-3A, Etymotic Research, Elk Grove Village, IL, USA).

The speech syllable [da] is a 170 ms voiced, six-formant stop consonant with a fundamental frequency (F0) of 100 Hz that was constructed using a Klatt-based synthesizer at 20 kHz (Figure 1A). During the consonant-vowel transition (i.e., the /d/ to /a/), the lower three formants linearly change (F1: 400-720 Hz, F2: 1700-1240 Hz, F3: 2580-2500 Hz), while the F0 and upper three formants remain steady (F0: 100 Hz, F4: 3300 Hz, F5: 3750 Hz, F6: 4900 Hz). During the vowel portion of the stimulus (i.e., the /a/), the F0 and six formants remain steady.

During the recording session, children sat in a recliner chair and watched a movie of their choice. To encourage compliance, the left ear was unoccluded so the child could hear the soundtrack of the movie (< 40 dB SPL in sound field). At least 4200 stimulus trials were presented to obtain 4000 artifact-free trials. Electrodes were placed at Cz for active (non-inverting), right and left earlobes for unlinked references (inverting), and ±1 cm on either side of Fpz for ground (CMS/DRL). All offsets were kept below 50 mV. Within the BioSEMI ActiABR module for LabView 2.0 (National Instruments, Austin, TX, USA), and per the hardware's limitation, responses were online filtered from 100-3000 Hz (6 dB/octave roll-off) and

digitized at 16.384 kHz. To open the high pass to 0.1 Hz, offline amplification of 6 dB/octave was performed in the frequency domain in MATLAB (The Mathworks, Inc., Natick, MA, USA) using custom programs. Responses were then bandpass filtered to the frequency region of interest (70-2000 Hz, Butterworth filter, 12 dB/octave roll off, zero phase shift), epoched from -40 to 210 ms (stimulus onset at 0 ms), baseline-corrected relative to the prestimulus period, and artifact rejected at ± 35 µV. To emphasize the stimulus envelope, responses to the alternating polarities were added (Aiken and Picton, 2008).

2.4 Frequency following response parameters
The FFR is an auditory evoked potential that offers insights into various aspects of neural sound processing (Krizman and Kraus, 2019; Skoe and Kraus, 2010). Four neural parameters of interest, *neural timing, spectral coding*, *response stability,* and *nonstimulus activity* were included in this investigation (Figure 1).

[Figure 1 here]

2.4.1 Neural timing
Neural timing was examined by measuring peak latencies of the FFR, which occur at periodic intervals that roughly correspond to the periodicity of the fundamental frequency (F0) and are thought to reflect phase locking to the stimulus (Krizman and Kraus, 2019; Skoe and Kraus, 2010). Latencies were identified in Neuroscan (Neuroscan Edit 4.5, Compumedics, Charlotte, NC) using a local maximum and minimum detection algorithm followed by manual verification using a blind procedure outlined in (Anderson et al., 2010a). Thirty eight peaks and troughs were chosen, and for statistical analyses, a composite timing measure was calculated by averaging the peak and trough latencies within the 20-160 ms response region.

2.4.2 Spectral coding
The scalp-recorded FFR robustly represents the fundamental frequency (F0) and harmonics of speech (Greenberg et al., 1987; Krishnan et al., 2004, 2005; Russo et al., 2004; Xu et al., 2006). To determine the frequency representation within the FFR, a fast Fourier transform (FFT) was applied to extract the spectral amplitudes of the fundamental frequency (F0) and its integer harmonics up to 1000 Hz within the 5-170 ms portion of the response. A 16,384 point FFT was computed with an 82.5 ms ramp, and amplitudes were calculated over 40 Hz bins centered at the F0 and integer harmonics. For statistical purposes, a composite measure was created by averaging all spectral amplitudes from the fundamental frequency (F0) through the tenth harmonic (H10).

2.4.3 Response stability
To measure response stability over the course of the recording session, first, two sub-averages of the FFR response were calculated. Each sub-average comprised 2000 sweeps, 1000 of each stimulus polarity, so that the odd epochs of each polarity were used to create one sub-average and the even epochs of each polarity were used to create the other sub-average (Hornickel and Kraus, 2013). The sub-averages were then correlated to compute a Pearson product-moment correlation coefficient (r). For statistical purposes, data were Fisher (z) transformed.

2.4.4 Nonstimulus activity
To quantify the non-evoked neural activity of the FFR, the root-mean-square (RMS) amplitude of the 40-ms interval preceding the stimulus was calculated.

2.5 Data monitoring and outlier checking
Prior to statistical analyses, data were examined for outliers. Particular care was taken to remove as few outliers as possible. Data points were only excluded if 1) there was a technical error in collection, or 2) if > 3 standard deviations (SD) above/below mean. Because these neurophysiological measures are distinct, an outlier for one neurophysiological measure was not necessarily an outlier for the other neurophysiological measures within that year. In addition, if a participant's data was excluded at one time point (e.g., at age 3), but had data from other time points within 3 SD (e.g., at ages 4 and 5), these values were included. In all, 10 out of 463 (2.16%) data points were removed for neural timing analyses, 16 out of 463 (3.46%) for spectral coding, 10 out of 463 (2.16%) for response stability, and 19 out of 463 (4.10%) for nonstimulus activity. Across all measures and all years (n = 463 test points), 26 subjects had one outlier datapoint, 4 subjects had two outlier datapoints, 2 subjects had three outlier datapoints, and 4 subjects had four outlier datapoints. 139 subjects had zero outlier datapoints. Please see Table 1C for more information regarding outlier exclusion by year of study.

For both spectral coding and nonstimulus activity, values were converted to nanovolts instead of microvolts due to small variance estimates. Structural equation models, like all methods derived from Generalized Linear Model (GLM), are invariant to linear transformations; results are the same regardless of metric unit (i.e., if analyzed in nanovolts, microvolts, etc.). Thus, results for these measures are reported in nanovolts instead of microvolts; to convert the growth curve model estimate values from nanovolts to microvolts, means can be divided by 1×10^3 and variance terms by 1×10^6.

2.6 Statistical analyses – growth curve modeling
For each neurophysiological measure (i.e., neural timing, spectral coding, response stability, and nonstimulus activity), a series of growth curve models were run to identify the functional form that best represented within- and between-individual changes (Please see Figure 2 for path diagram reflecting the structure of growth curve modeling). The series of growth curve models were run through a multi-level structural equation model where each individual's age at test (e.g., Age3 = 3.57 years) was used as the basis coefficient for estimating change over time. In other words, each participant's precise age at test (e.g., 3.47) is loaded into the model as their "age" for the year they visited the lab, and predictions in the neural measure are made for that exact age. In addition to accounting for this variance in age, the growth curve models also account for variance in time between an individual's test points (e.g., 12 months to 20 months, etc). Included in the models were estimates of intercept (I) and slope (S) means and variances. Constraining intercept values to 1 allowed us to determine the neural measure's initial status (i.e. at age 3), while constraining slope values to age of test allowed us to determine linear changes in growth over time. Model comparisons were performed using Chi-Square Difference Tests based on log likelihood values and scaling correction factors obtained within the MLR estimator (Satorra and Bentler, 2010). All analyses were run in Mplus (Muthén and Muthén, 2017). P values less than .05 were considered significant.

[Figure 2 here]

The series of growth curve models were implemented as follows. First, we ran a full growth curve model to examine individual differences within our dataset, as well as individual differences with respect to development (i.e., rates of change). In this model (Model 1), all parameters were freely estimated: intercept mean, slope mean, intercept variance, intercept slope, intercept and slope mean covariance, and residual variances. Next, we ran a simplification of Model 1 to estimate the overall change in the respective neurophysiological measure over time. In this model (i.e., Model 2), fixing slope variance to zero allowed us to test the null hypothesis that all individuals develop at the same rate. Finally, we ran an even simpler model, Model 3, by fixing slope mean and slope variance to zero to test the null hypothesis that there are no changes in neurophysiological development over time.

8

To answer our research questions we systematically excluded growth curve parameters in the series of models (M1, M2, M3) and then performed model comparisons (M1 vs M2, M2 vs. M3). Comparing models M2 vs M3 allowed us to determine whether or not improvement in fit was attributable to overall change over time (i.e., growth curve parameter: slope mean), while comparing M2 vs M1 allowed us to determine whether the improvement in fit is attributable to individual differences in rates of change (i.e., growth curve parameter: slope variance).

2.7 Measures of model fit
To determine standardized measures of model fit, the series of growth models were run a second time through a traditional structural equation model where fixed ages were used (e.g., Age3 = 3.0, for all individuals) as the basis coefficients for estimating change over time. These indices included the Root Mean Square Error of Approximation (RMSEA), the Comparative Fit Index (CFI), and the Tucker-Lewis index (TLI). For a description of these indices please see (Hooper et al., 2008). For RMSEA, values less than 0.04 are considered to be an "excellent" fit, less than 0.07 a "good" fit, and less than 0.1 a "fair" fit (Steiger, 2007); for CFI and TLI, models with an "excellent" fit are greater than 0.95 (Hu and Bentler, 1999).

By running the growth curves through the multi-level structural equation model (with age at test as the basis coefficient) and traditional structural equation model (with fixed age as the basis coefficient) we can 1) determine the fit of the models with greater age specificity to determine precise estimates, and 2) compare the models using standardized and widely-used "goodness of fit" indices.

2.8 Statistical analyses – other
For each neurophysiological measure, we examined the Intraclass Correlation Coefficient (ICC), ρ, which reflects the proportion of variance in the outcome variable explained by each model's total variance. In other words, the ICC allows us to calculate the amount of variation unexplained by model predictors and how that relates to overall unexplained variance. In addition, to evaluate the collinearity of the neurophysiological measures, we performed a series of Pearson Correlations controlling for age. Finally, to investigate the influence of sex on auditory neurophysiological development, growth curve analyses were run including sex as a time-invariant factor. Because sex was not a significant predictor of growth for the four neurophysiological measures, these results are excluded from the main text and reported in the Supplementary Appendix. Of note, including sex as a factor revealed overall sex differences in each of the four neurophysiological measures, in that females had FFRs that were faster, more robust in representing spectral information, more consistent, and less noisy compared to their male peers. As sex differences are not the main focus of this paper, results are provided in the Supplementary Appendix as well. For a recent report of developmental sex differences in subcortical auditory processing, please see (Krizman et al., 2019).

3. Results
3.1 Summary
Fit statistics and parameter estimates for the series of growth curve models are provided in Table 2 and Table 3, respectively. Estimates for the best fitting growth curve model are reported; for all growth curve parameter estimates, please see Supplementary Appendix Table A.1. Fit statistics for neural timing, spectral coding, and response stability growth curve models reflected appropriately fitting models. For neural timing and response stability, the second model (M2) fit best, suggesting these neurophysiological measures improved over time: as children aged, neural responses became earlier and more consistent. For spectral coding, the first model (M1) fit best, suggesting frequency representation also improves over time (spectral coding became stronger with age), and that there were individual differences in rates of change. We observed relatively high Intraclass Correlation Coefficient (ICC) values, ρ, for neural timing (ρ = 0.731) and spectral coding (ρ = 0.670), a moderately-sized ICC value for response stability (ρ = 0.426),

and a relatively low ICC value for nonstimulus activity (ρ = 0.294). Given high ICC values indicate high similarity between values from the same measure, and low ICC values reflect little to no similarity between values from the same measure, these results suggest that neural timing, spectral coding, and response stability are reliable indices of their respective facet of sound processing, and therefore, how sound processing changes over time.

[Table 2 & 3 here]

3.2 Neural timing development

Model comparisons of the latency growth curves show M2 had the best fit (Table 2). By testing whether or not the addition of the mean slope parameter significantly improved the growth curve (M2 vs M3), we found the overall model fit improved (Chi square difference test: M2 vs M3, p < 0.01) and had an overall "fair" fit based on traditional "goodness of fit" indices (M2: RMSEA = 0.103, CFI = 0.847, TLI = 0.903). In contrast, the addition of the slope variance parameter (M1) did not significantly improve the model (Chi square difference test: M1 vs M2, p = 0.285).

The model comparisons for Model 2 affirms two of our three hypotheses with respect to neural timing: 1) auditory encoding of neural timing changes across childhood (Mean slope: Estimate = -0.018, SE = 0.005, p < 0.001) and 2) individual differences exist in neural timing (Intercept variance: Estimate = 0.037, SE = 0.005, p < 0.001). Because freely estimating slope variance did not improve the model (M1 vs M2), our dataset does not support the notion that there are individual differences in rates of change in neural timing.

Model 2's parameter estimates (Table 3) indicate the average latency value of the neural response was 90.744 milliseconds at age 3 (Mean intercept: Estimate = 90.744, SE = 0.030, p < 0.001) and became significantly earlier each year by 0.018 milliseconds. Please see Figure 3 for graphical representation of neural timing development.

[Figure 3 here]

3.3 Spectral coding development

Model comparisons of the spectral coding growth curves show M1 had the best fit. The addition of the mean slope parameter significantly improved the growth curve (Chi square difference test: M2 vs M3, p < 0.04), as did the addition of the slope variance parameter (Chi square difference test: M1 vs M2, p = 0.006). Model 1 had an "excellent" fit based on traditional "goodness of fit" indices (M1: RMSEA = 0.04, CFI = 0.970, TLI = 0.979).

The model comparisons affirms our three hypotheses with respect to spectral coding: 1) auditory spectral coding changes across childhood (Mean slope: Estimate = 0.318, SE = 0.151, p = 0.035), 2) individual differences in spectral coding exist (Intercept variance: Estimate = 44.414, SE = 11.633, p < 0.001), and 3) individual differences in the development of spectral coding exist (Slope variance: Estimate = 0.691, SE = 11.633, p < 0.001). Each individual's spectral coding starting level is related to developmental changes over time (latent variable covariances of intercept & slope: Estimate = -4.167, SE = 1.796, p = 0.020), such that lower spectral coding values at age 3 are associated with greater change in spectral coding over time each year.

Model 1's parameter estimates indicate the average level of spectral coding was 0.017 microvolts at age 3 (Mean intercept: Estimate = 17.463, SE = 0.848, p < 0.001) and became significantly larger each year by 0.003 microvolts. Please see Figure 4 for graphical representation of spectral coding development.

[Figure 4 here]

3.4 Response stability development

Model comparisons of the response stability growth curves show M2 had the best fit. By testing whether or not the addition of the mean slope parameter significantly improved the growth curve (M2 vs M3), we found the overall model fit improved (Chi square difference test: M2 vs M3, p = 0.03) and was "excellent" based on traditional "goodness of fit" indices (M2: RMSEA = 0.00, CFI = 1.00, TLI = 1.055). In contrast, the addition of the slope variance parameter (M1) did not significantly improve the model (Chi square difference test: M1 vs M2, p = 0.178).

Similar to neural timing, the model comparisons affirms two of our three hypotheses with respect to response stability: 1) auditory response stability changes across childhood development (Mean slope: Estimate = 0.016, SE = 0.008, p = 0.043) and 2) there are individual differences in response stability overall (Intercept variance: Estimate = 0.025, SE = 0.004, p < 0.001). Because freely estimating slope variance did not improve the model (M1 vs M2), our dataset does not support the notion that there are individual differences in response stability rates of change.

Model 2's parameter estimates indicate the average response stability value was 0.597 at age 3 (Mean intercept: Estimate = 0.597, SE = 0.045, p < 0.001) and became significantly larger each year by 0.016 units. Please see Figure 5 for graphical representation of response stability development.

[Figure 5 here]

3.5 Nonstimulus activity

Model comparisons of the nonstimulus activity growth curves show M1 had the best fit, despite being a poorly fitting model overall (M1: RMSEA = 0.066, CFI = 0.576, TLI = 0.703).

[Table 4 here]

3.6 Relationships between neural measures

To determine collinearity between neural measures, we ran a series of Pearson's correlations (Table 4). Controlling for age, weak relationships were found between nonstimulus activity and neural timing (r = -0.086), neural timing and response stability (r = -0.217), and nonstimulus activity and spectral coding (r = 0.158). Moderate relationships were observed between neural timing and spectral coding (r = -0.400), spectral coding and response stability (r = 0.656), and nonstimulus activity and response stability (r = -0.446). No strong correlations (r > .7) were evident among the neural measures.

4. Discussion

4.1 Auditory neurophysiological development evident in childhood

Using the novel application of growth curve modeling, we show that the neural processing of speech continues to mature within individuals over the course of childhood: neurophysiological responses became earlier, more robust, and more stable. In addition to providing longitudinal evidence of central auditory system development throughout this age range, these results demonstrate there are individual differences in neural processing of various sound ingredients, and that for some of these ingredients (i.e., spectral coding), there are individual differences in how neural development unfolds over time.

These findings align with cross-sectional evidence showing neurophysiological development of the auditory system. Using the FFR, age-related differences in neural processing of timing and frequency information have been observed in infants, 1 to 3 months of age (Jeng et al., 2010) and 3 to 10 months of age (Anderson et al., 2015), as well as children, ages 3-5 and 8-12 years (Johnson et al., 2008). In addition, changes in timing and spectral coding, as well as response stability and nonstimulus activity occur across the lifespan (Krizman et al., 2019; Skoe et al., 2015). These studies show evidence of multiple maturational timelines for measures of subcortical processing, and that development of these measures continues through late childhood (ages 3-8). Through our use of a different speech stimulus and a large longitudinal sample, the present findings reinforce these cross-sectional observations, yet for the first time show developmental changes are evident between- *and* within-individuals.

4.2 Auditory neurophysiological changes parallel perceptual development
These findings also align with psychophysical evidence showing gradual auditory perceptual development over the first decade of life (Sanes and Woolley, 2011). For example, temporal integration, one aspect of temporal processing and the process in which information is summed over time, develops through age 6 when tested with experimental paradigms of duration discrimination (Jensen and Neff, 1993), gap detection (Wightman et al., 1989), and the precedence effect (Litovsky, 1997). Here, we examined temporal processing through a measure of neural timing, and found that as children age, processing of sound becomes faster (i.e., earlier latencies). Nervous system timing (latency) is inversely related to white matter density (Eggermont and Moore, 2012), and latency changes are likely to reflect a rapid increase of axonal myelination throughout the auditory system (Moore and Linthicum, 2007). Given the parallel between these two lines of research, and the objectivity of the FFR approach, our results corroborate the psychophysical evidence. Importantly, we show that sensory processing develops without relying on behavioral tests whose interpretation can be complicated by cognitive influences, such as attention.

Also in line with the psychophysical literature is our finding that spectral coding changes over the course of early childhood. In terms of auditory perceptual indices of spectral processing, frequency resolution matures early on, reaching adult-like levels by 6 months of age (Spetner and Olsho, 1990). In contrast, more complex spectral processing matures later in childhood. For example, frequency discrimination, a task that requires one to detect differences in frequency presented successively, matures around age 10 for low frequency tones (Maxon and Hochberg, 1982). Tasks requiring detection of frequency and amplitude modulations, cues that are important for speech perception, follow a relatively prolonged timeline, such that maturation continues beyond age 12 (Banai et al., 2011). Task-related differences in maturational trajectories illustrate how some perceptual tests draw on fundamental cochlear mechanisms (i.e., "hearing") that matures early in life, while others involve auditory and non-auditory centers that follow a longer developmental course. Our finding that spectral coding develops throughout childhood suggests that this measure could be an index of complex sound processing that offers objectivity into the maturation of frequency representation.

While neural timing and spectral coding have perceptual analogs, response stability does not. Response stability is thought to be an index of how replicable the nervous system represents a stimulus over time (Hornickel and Kraus, 2013; Krizman and Kraus, 2019), reflecting both the brain's endurance and reliability within a testing session. Although no behavioral indices exist to draw a direct comparison, theoretically, response stability might inform our knowledge of "internal noise", or within-individual variability. In Signal Detection Theory, one assumption is that "internal noise" can increase variance in perceptual indices (MacMillan, 2002). For example, infants are considered "broadband listeners" with more internal noise; over time, internal noise decreases and infants tune in to specific sound features rather than rely on a Gestalt representation (Saffran et al., 2007; Werner, 1996; Werner et al., 2012). After infancy, internal noise in the auditory system continues to decrease from childhood to adulthood across a range of tasks, including intensity discrimination (Buss et al., 2006, 2009) and detections of tones in noise

(Allen and Wightman, 1994). In the present study, we see responses become more consistent over time. This finding could reflect 1) the gradual dissipation of internal noise within childhood, 2) that neural encoding of stimulus features is more equivalent across trials, or 3) both. Future research is needed to further understand this neurophysiological parameter, and how developmental changes in response stability may be tied to changes in behavior (e.g., reading disorders (Hornickel and Kraus, 2013)). Furthermore, while response stability may grossly reflect "internal noise", the possibility of other noise sources influencing this measure remains. Future research is necessary to disentangle the contribution of neural from non-neural noise of this parameter specifically, as well as the other measures examined in this study.

The growth curve model examining the development of nonstimulus activity fit poorly, perhaps due to a few reasons. First, a "poor fit" of a growth curve might be driven by noisy data, where between- and within-individual changes are too variable to be modeled statistically. Such a result could arise because the measure of interest reflects both internal and external noise. Indeed, the ICC for nonstimulus activity was relatively weak ($\rho = 0.294$), suggesting this neurophysiological measure reflects more variability and/or noise in any given observation relative to the other neurophysiological measures. One other possible contributor to a poorly fitting model is multiphasic growth (i.e., more than one growth phase). Unfortunately, to determine whether or not a sample follows a multiphasic growth pattern requires a large number of subjects, one which exceeds the present sample size. Future research should examine how nonstimulus activity relates to general indices of neural and non-neural noise and should extend the sample size to a number sufficient to determine if nonstimulus activity development follows a multiphasic—rather than linear–growth trajectory.

4.3 Investigating neurophysiological development using growth curve modeling
For many years, neurophysiologic development of the auditory system was assessed using the auditory brainstem response, which is believed to be adult-like by about 18 months (Jerger and Hall, 1980). This led to the assumption that the auditory system was stable by the second year of life. However, many investigations in the past few decades have revisited the notion of auditory system development, revealing maturation through at least young adulthood (Krizman et al., 2019; Ponton et al., 2000; Skoe et al., 2015). While these studies made considerable headway in redefining the notion of auditory system development, they did so using a cross-sectional approach, leaving the possibility that between-individual changes differ from within-individual changes. Few examinations have adopted a longitudinal approach to understand within-individual changes in the auditory system, and whether or not there are individual differences in rates of change. In the present study, we use a large sample to provide evidence that reinforces these cross-sectional examinations, yet shows for the first time that auditory changes are evident within individuals.

Neurophysiological data were analyzed using growth curve modeling, a statistical approach that is burgeoning in the fields of social, psychological, and behavioral sciences due to advancements in technology and computing. Although the benefits of growth curve modeling are many, the number of neurophysiological investigations employing this approach are few. Our novel application of this statistical approach merits a discussion of its advantages and limitations.

One advantage of growth curve modeling is that it permits the inclusion of all data points collected on an individual. This is in contrast to traditional statistical approaches, such as repeated measures analysis of variance, which require a complete dataset for included individuals and therefore excludes those that do not meet this requirement. In growth curve modeling, each data point adds value to a model, even if data was collected at only one time point (i.e., cross-sectional) or a subset of those time points (such as years 1, 2, and 3 in a 4-year study). Though technically not longitudinal data, individuals with one time point of data add value to the growth curve model by informing the distribution of that individual's age range that would otherwise not be known. However, while these "cross-sectional" data points add value to a growth

curve, it is important to note that the relative value is small in comparison to individuals with multiple time points of data collected; in our case, those with five years' worth of data added the most value to the model.

A second advantage of growth curve modeling is the ability to characterize the functional form of developmental changes, such as linear or non-linear trends. Here, we see that the growth in our longitudinal data was linear, rather than cubic or quadratic. In other words, developmental changes occurred steadily and gradually over time. This result is novel, especially considering the non-linear developmental path of many skills and behaviors (e.g., language acquisition) (Bates et al., 1995). Though we do not suspect linear changes to continue throughout the lifespan (cf. Skoe et al. 2015) it is interesting to think of this age range as a linear developmental period.

A third and final advantage of growth curve modeling is the statistical determination of individual differences and individual differences in rates of change. Here, we see that there are individual differences across the parameters, suggesting auditory processing of neural timing, spectral coding, and response stability are likely explained by differences in genetics, environmental factors, etc. Moreover, individual differences with respect to change were evident in spectral coding development, meaning there is more than one maturational path in this age range. In contrast, we did not see individual differences in rates of change for neural timing or response stability measures. This is not to say that individual differences do not exist; rather, our dataset does not support this result. Future studies are needed to investigate differential rates of change with respect to neural timing and response stability in childhood auditory development.

One drawback of growth curve modeling is the necessity of a large sample, which for developmental investigations might prove unrealistic. While we provide results that were sufficiently powered, a greater number of individuals would be needed to investigate how, and why children develop at distinct rates. To address the question of whether or not the neural measures reflect distinct processes in our data, Pearson correlations were run to determine collinearity. No strong correlations ($r > .9$) were observed between the neurophysiological measures at any age group, suggesting these indices reflect distinct features of auditory nervous system function.

4.4 Links between auditory development & communication skills
Throughout childhood, the auditory system plays a large role in the acquisition of many communication skills. For example, reading involves mapping sounds to meaning, while speech perception requires parsing relevant from irrelevant information within an incoming auditory stream. Strength of auditory-neurophysiological sound processing is linked to a number of these communication skills. For example, enhancements in temporal and spectral coding are related to school-aged children's ability to hear in noisy environments (Anderson et al., 2010b, 2010a). For preschool-aged children, development of the fundamental frequency of speech (one frequency included in our spectral coding measure) tracks with improvements in hearing in noise (Thompson et al., 2017). Literacy and reading competencies are linked to neural stability (Centanni et al., 2014; Hornickel and Kraus, 2013), auditory system timing (Ahissar et al., 2000; Banai et al., 2009) and processing of detailed acoustic features such as consonants (Kraus et al., 1996; Tallal, 1980).

To understand the relationship between neural processing of sound and clinical disorders, a critical first step is to determine how these auditory neurophysiological measures mature in a typically developing population, with the ultimate goal to improve clinical diagnostics and interventions. Given the links between auditory function and communication, an objective approach like the FFR, could provide a more reliable method for identification and diagnosis of developmental disorders related to reading and language, including developmental dyslexia and specific language impairment. For example, as an objective index of distinct facets of auditory maturation, the FFR could reveal aberrant patterns of

14

development that suggest delays earlier. Future studies should examine deviations in auditory development in clinical populations, and/or how abnormal auditory system development could contribute to disordered sound processing.

4.5 Future directions

Participants in this study were monolingual English speakers from similar socioeconomic backgrounds, and therefore reflect a small subset of the broader population. Future research should investigate the influence of second language experience and socioeconomic standing on the neurophysiological development of the auditory system. This is especially important as previous research shows spectral coding, response stability, and nonstimulus activity vary along socioeconomic dimensions in adolescents (Skoe et al., 2013) and based on language experience (Krizman et al., 2012; Skoe et al., 2017). Second, future research should investigate how performance on language or auditory tests track with objective auditory system development, as indexed by the FFR, to determine behavioral consequences of neurophysiological maturation. A third line of future research is the investigation of auditory maturation prior to 3 and 4 years of age. While a few studies have demonstrated auditory system development evolves over the first two years of life, a gap in the literature between 10 months and 3 years of age remains. A systematic, longitudinal examination is needed to confirm these maturational changes and to fully understand auditory development from the onset.

5. Conclusion

Through a large longitudinal dataset analyzed through growth curve modeling, we observe the sharpening of neurophysiological processing of sound within individuals in childhood: neural timing became faster, spectral coding stronger, and response stability more consistent over time. These findings align with the perceptual and cross-sectional investigations of auditory development, and show that auditory maturation extends into late childhood using an objective and longitudinal approach.

Declaration of interest

None.

References

Abdala, C., and Keefe, D.H. (2006). Effects of middle-ear immaturity on distortion product otoacoustic emission suppression tuning in infant ears. J. Acoust. Soc. Am. *120*, 3832–3842.

Ahissar, M., Protopapas, A., Reid, M., and Merzenich, M.M. (2000). Auditory processing parallels reading abilities in adults. Proc. Natl. Acad. Sci. U. S. A. *97*, 6832–6837.

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

Allen, P., and Wightman, F. (1994). Psychometric Functions for Children's Detection of Tones in Noise. J. Speech Lang. Hear. Res. *37*, 205.

Anderson, S., Skoe, E., Chandrasekaran, B., and Kraus, N. (2010a). Neural Timing Is Linked to Speech Perception in Noise. J. Neurosci. *30*, 4922–4926.

Anderson, S., Skoe, E., Chandrasekaran, B., Zecker, S., and Kraus, N. (2010b). Brainstem correlates of speech-in-noise perception in children. Hear. Res. *270*, 151–157.

Anderson, S., Parbery-Clark, A., White-Schwoch, T., and Kraus, N. (2012). Aging Affects Neural Precision of Speech Encoding. J. Neurosci. *32*, 14156–14164.

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

Banai, K., Hornickel, J., Skoe, E., Nicol, T., Zecker, S., and Kraus, N. (2009). Reading and Subcortical Auditory Function. Cereb. Cortex *19*, 2699–2707.

Banai, K., Sabin, A.T., and Wright, B.A. (2011). Separable developmental trajectories for the abilities to detect auditory amplitude and frequency modulation. Hear. Res. *280*, 219–227.

Bates, E., Dale, P.S., and Thal, D. (1995). Individual differences and their implications for theories of language development. In: P. Fletcher, B. MacWhinney, eds. Handbook of Child Language. Oxford: Basil Blackwell.

Bidelman, G.M. (2015). Multichannel recordings of the human brainstem frequency-following response: Scalp topography, source generators, and distinctions from the transient ABR. Hear. Res. *323*, 68–80.

Buss, E., Hall, J.W., and Grose, J.H. (2006). Development and the role of internal noise in detection and discrimination thresholds with narrow band stimuli. J. Acoust. Soc. Am. *120*, 2777–2788.

Buss, E., Hall, J.W., and Grose, J.H. (2009). Psychometric functions for pure tone intensity discrimination: Slope differences in school-aged children and adults. J. Acoust. Soc. Am. *125*, 1050–1058.

Carré, R., Divenyi, P., and Mrayati, M. (2017). Speech: A dynamic process. Berlin, Boston: De Gruyter.

Centanni, T.M., Booker, A.B., Sloan, A.M., Chen, F., Maher, B.J., Carraway, R.S., Khodaparast, N., Rennaker, R., LoTurco, J.J., and Kilgard, M.P. (2014). Knockdown of the Dyslexia-Associated Gene

Kiaa0319 Impairs Temporal Responses to Speech Stimuli in Rat Primary Auditory Cortex. Cereb. Cortex *24*, 1753–1766.

Chandrasekaran, B., and Kraus, N. (2010). The scalp-recorded brainstem response to speech: Neural origins and plasticity. Psychophysiology *47*, 236–246.

Eggermont, J.J., and Moore, J.K. (2012). Morphological and Functional Development of the Auditory Nervous System. In: L. Werner, R.R. Fay, A.N. Popper, eds. Human Auditory Development. New York, NY: Springer New York, pp. 61–105.

Eggermont, J.J., Brown, D.K., Ponton, C.W., and Kimberley, B.P. (1996). Comparison of Distortion Product Otoacoustic Emission (DPOAE) and Auditory Brain Stem Response (ABR) Traveling Wave Delay Measurements Suggests Frequency-Specific Synapse Maturation: Ear Hear. *17*, 386–394.

Frick, R.W. (1985). Communicating emotion: The role of prosodic features. Psychol. Bull. *97*, 412–429.

Greenberg, S., Marsh, J.T., Brown, W.S., and Smith, J.C. (1987). Neural temporal coding of low pitch. I. Human frequency-following responses to complex tones. Hear. Res. *25*, 91–114.

Hartley, D.E.H., Wright, B.A., Hogan, S.C., and Moore, D.R. (2000). Age-Related Improvements in Auditory Backward and Simultaneous Masking in 6- to 10-Year-Old Children. J. Speech Lang. Hear. Res. *43*, 1402.

Hooper, D., Coughlan, J., and Mullen, M.R. (2008). Structural Equation Modelling: Guidelines for Determining Model Fit. Electron. J. Bus. Res. Methods *6*, 53–60.

Hornickel, J., and Kraus, N. (2013). Unstable Representation of Sound: A Biological Marker of Dyslexia. J. Neurosci. *33*, 3500–3504.

Hu, L., and Bentler, P.M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. Struct. Equ. Modeling *6*, 1–55.

Jeng, F.-C., Schnabel, E.A., Dickman, B.M., Hu, J., Li, X., Lin, C.-D., and Chung, H.-K. (2010). Early Maturation of Frequency-Following Responses to Voice Pitch in Infants with Normal Hearing. Percept. Mot. Skills *111*, 765–784.

Jensen, J.K., and Neff, D.L. (1993). Development of Basic Auditory Discrimination in Preschool Children. Psychol. Sci. *4*, 104–107.

Jerger, J., and Hall, J. (1980). Effects of Age and Sex on Auditory Brainstem Response. Arch. Otolaryngol. *106*, 387–391.

Johnson, K.L., Nicol, T., Zecker, S.G., and Kraus, N. (2008). Developmental Plasticity in the Human Auditory Brainstem. J. Neurosci. *28*, 4000–4007.

Kraus, N., McGee, T.J., Carrell, T.D., Zecker, S.G., Nicol, T.G., and Koch, D.B. (1996). Auditory neurophysiologic responses and discrimination deficits in children with learning problems. Science *273*, 971–973.

Krishnan, A., Xu, Y., Gandour, J.T., and Cariani, P.A. (2004). Human frequency-following response: representation of pitch contours in Chinese tones. Hear. Res. *189*, 1–12.

Krishnan, A., Xu, Y., Gandour, J., and Cariani, P. (2005). Encoding of pitch in the human brainstem is sensitive to language experience. Cogn. Brain Res. *25*, 161–168.

Krizman, J., and Kraus, N. (2019). Analyzing the FFR: A tutorial for decoding the richness of auditory function. Hear. Res. *382*, 107779.

Krizman, J., Marian, V., Shook, A., Skoe, E., and Kraus, N. (2012). Subcortical encoding of sound is enhanced in bilinguals and relates to executive function advantages. Proc. Natl. Acad. Sci. *109*, 7877–7881.

Krizman, J., Bonacina, S., and Kraus, N. (2019). Sex differences in subcortical auditory processing emerge across development. Hear. Res. *380*, 166–174.

Krizman, J., Lindley, T., Bonacina, S., Colegrove, D., White-Schwoch, T., and Kraus, N. (2020). Play Sports for a Quieter Brain: Evidence From Division I Collegiate Athletes. Sports Health *12*, 154–158.

Lavigne-Rebillard, M., and Pujol, R. (1987). Surface aspects of the developing human organ of Corti. Acta Oto-Laryngol. Suppl. *436*, 43–50.

Litovsky, R.Y. (1997). Developmental changes in the precedence effect: Estimates of minimum audible angle. J. Acoust. Soc. Am. *102*, 1739–1745.

Macias, M.M., and Twyman, K.A. (2011). Developmental and behavioral pediatrics. Chapter 11: Speech and Language Development and Disorders. Elk Grove Village, IL: American Academy of Pediatrics.

MacMillan, N.A. (2002). Signal Detection Theory. In: H. Pashler, ed. Stevens' Handbook of Experimental Psychology. Hoboken, NJ, USA: John Wiley & Sons, Inc.

Maxon, A.B., and Hochberg, I. (1982). Development of psychoacoustic behavior: sensitivity and discrimination. Ear Hear. *3*, 301–308.

Moore, D.R. (2002). Auditory development and the role of experience. Br. Med. Bull. *63*, 171–181.

Moore, J.K., and Linthicum, F.H. (2007). The human auditory system: A timeline of development. Int. J. Audiol. *46*, 460–478.

Moore, D.R., Ferguson, M.A., Edmondson-Jones, A.M., Ratib, S., and Riley, A. (2010). Nature of Auditory Processing Disorder in Children. PEDIATRICS *126*, e382–e390.

Muthén, L.K., and Muthén, B.O. (2017). Mplus User's Guide. Eighth Edition.

Otto-Meyer, S., Krizman, J., White-Schwoch, T., and Kraus, N. (2018). Children with autism spectrum disorder have unstable neural responses to sound. Exp. Brain Res. *236*, 733–743.

Ponton, C.W., Eggermont, J.J., Coupland, S.G., and Winkelaar, R. (1992). Frequency-specific maturation of the eighth nerve and brain-stem auditory pathway: Evidence from derived auditory brain-stem responses (ABRs). J. Acoust. Soc. Am. *91*, 1576–1586.

Ponton, C.W., Eggermont, J.J., Kwong, B., and Don, M. (2000). Maturation of human central auditory system activity: evidence from multi-channel evoked potentials. Clin. Neurophysiol. *111*, 220–236.

18

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

Saffran, J.R., Werker, J.F., and Werner, L.A. (2007). The Infant's Auditory World: Hearing, Speech, and the Beginnings of Language. In: W. Damon, R.M. Lerner, eds. Handbook of Child Psychology. Hoboken, NJ, USA: John Wiley & Sons, Inc.

Sanes, D.H., and Woolley, S.M.N. (2011). A Behavioral Framework to Guide Research on Central Auditory Development and Plasticity. Neuron *72*, 912–929.

Satorra, A., and Bentler, P.M. (2010). Ensuring Positiveness of the Scaled Difference Chi-square Test Statistic. Psychometrika *75*, 243–248.

Sharma, A., Kraus, N., J. McGee, T., and Nicol, T.G. (1997). Developmental changes in P1 and N1 central auditory responses elicited by consonant-vowel syllables. Electroencephalogr. Clin. Neurophysiol. Potentials Sect. *104*, 540–545.

Skoe, E., and Kraus, N. (2010). Auditory Brain Stem Response to Complex Sounds: A Tutorial: Ear Hear. *31*, 302–324.

Skoe, E., Krizman, J., and Kraus, N. (2013). The Impoverished Brain: Disparities in Maternal Education Affect the Neural Response to Sound. J. Neurosci. *33*, 17221–17231.

Skoe, E., Krizman, J., Anderson, S., and Kraus, N. (2015). Stability and Plasticity of Auditory Brainstem Function Across the Lifespan. Cereb. Cortex *25*, 1415–1426.

Skoe, E., Brody, L., and Theodore, R.M. (2017). Reading ability reflects individual differences in auditory brainstem function, even into adulthood. Brain Lang. *164*, 25–31.

Spetner, N.B., and Olsho, L.W. (1990). Auditory frequency resolution in human infancy. Child Dev. *61*, 632–652.

Starr, A., Amlie, R.N., Martin, W.H., and Sanders, S. (1977). Development of auditory function in newborn infants revealed by auditory brainstem potentials. Pediatrics *60*, 831–839.

Steiger, J.H. (2007). Understanding the limitations of global fit assessment in structural equation modeling. Personal. Individ. Differ. *42*, 893–898.

Tallal, P. (1980). Auditory temporal perception, phonics, and reading disabilities in children. Brain Lang. *9*, 182–198.

Templin, M.C. (1957). Certain Language Skills in Children: Their Development and Interrelationships. University of Minnesota Press.

Thompson, E.C., Woodruff Carr, K., White-Schwoch, T., Otto-Meyer, S., and Kraus, N. (2017). Individual differences in speech-in-noise perception parallel neural speech processing and attention in preschoolers. Hear. Res. *344*, 148–157.

Werner, L.A. (1996). The Development of Auditory Behavior (or What the Anatomists and Physiologists Have to Explain). Ear Hear. *17*, 438–446.

Werner, L.A., Fay, R.R., and Popper, A.N. (2012). Human auditory development. New York; London: Springer.

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

White-Schwoch, T., Nicol, T., Warrier, C.M., Abrams, D.A., and Kraus, N. (2017). Individual Differences in Human Auditory Processing: Insights From Single-Trial Auditory Midbrain Activity in an Animal Model. Cereb. Cortex *27*, 5095–5115.

Wightman, F., Allen, P., Dolan, T., Kistler, D., and Jamieson, D. (1989). Temporal Resolution in Children. Child Dev. *60*, 611.

Wunderlich, J.L., and Cone-Wesson, B.K. (2006). Maturation of CAEP in infants and children: A review. Hear. Res. *212*, 212–223.

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

Tables and Figures

Table 1. A) Participant descriptive statistics broken down by age at test and year of test of the longitudinal study. B) Descriptive statistics for test-retest duration for each year of longitudinal study. C) Number of outliers for each neurophysiological measure that were excluded from growth curve analyses. Please see *Methods* for more information.

A) Participant descriptive statistics							
Age at test	**3yo**	**4yo**	**5yo**	**6yo**	**7yo**	**8+**	
Count (n)	82	141	95	70	48	27	
Average age (years)	3.47	4.49	5.51	6.52	7.49	8.54	
SD (years)	0.27	0.29	0.29	0.30	0.29	0.40	
Min age (years)	3.01	4.01	5.02	6.01	7.01	8.01	
Max age (years)	3.94	4.97	6.00	6.96	7.95	9.29	
Females (n)	42	63	40	30	22	11	
Males (n)	40	78	55	40	26	16	
Year of test	**Year 1**	**Year 2**	**Year 3**	**Year 4**	**Year 5**		
Count	175	115	85	62	26		
Females	81	51	36	28	12		
Males	94	64	49	34	14		
Average age (years)	4.02	5.12	6.21	7.32	8.36		
SD	0.62	0.63	0.65	0.65	0.54		
B) Test-retest duration (months)	**Year 1-2**	**Year 2-3**	**Year 3-4**	**Year 4-5**			
Average	12.90	12.77	12.45	11.39			
SD	2.00	1.78	1.24	1.66			
Min	7.79	10.64	11.04	6.67			
Max	20.99	22.14	18.33	13.50			
Median	12.25	12.29	12.12	11.86			
C) Outlier count (n)	**Year 1**	**Year 2**	**Year 3**	**Year 4**	**Year 5**	**All years**	**% of total sample**
Neural timing	2	2	3	3	0	10	2.16
Spectral coding	4	1	4	3	4	16	3.46
Response stability	7	0	1	2	0	10	2.16
Nonstimulus activity	8	4	3	2	2	19	4.10

Table 2. Fit statistics for the series of growth models for each neurophysiological measure. Shaded cells reflect fit statistics from the traditional structural equation model where the basis coefficients for estimating change over time were fixed (e.g., 3.0 for all 3-year-olds). Measures of model fit included Root Mean Square Error of Approximation (RMSEA), Comparative Fit Index (CFI), and Tucker-Lewis index (TLI). Unshaded cells represent model comparisons from the multi-level structural equation model where the basis coefficients for estimating change over time corresponded to age at test (e.g., 3.15 years, 3.87, etc.). The multi-level structural equation model comparisons (M1 vs M2 and M2 vs M3) were performed using Chi-Square Difference Tests based on log likelihood values (H0) and scaling correction factors obtained within the MLR estimator (SCFMLR), and were computed based on the difference test scaling correction (cd) and the Chi-Square difference test (TRd) values. Please see *Methods* for more information.

	M1	M2	M3
NEURAL TIMING			
Parameters	6	4	3
RMSEA	0.109	0.103	0.116
CFI	0.844	0.847	0.796
TLI	0.891	0.903	0.876
--2LL	154.367	153.759	146.458
H0	154.298	152.649	146.093
SCFMLR	1.332	1.340	1.545
		M1 vs M2	*M2 vs M3*
cd		1.316	0.726
TRd		2.507	18.071
pvalue		0.285	0.000
SPECTRAL CODING			
Parameters	6.000	4.000	3.000
RMSEA	0.040	0.054	0.061
CFI	0.970	0.941	0.922
TLI	0.979	0.962	0.953
--2LL	-1286.537	-1290.387	-1292.667
H0	-1286.508	-1290.124	-1292.666
SCFMLR	1.071	1.250	1.232
		M1 vs M2	*M2 vs M3*
cd		0.715	1.301
TRd		10.112	3.907
pvalue		0.006	0.048
RESPONSE STABILITY			
Parameters	6.000	4.000	3.000
RMSEA	0.000	0.000	0.000

		M1 vs M2	M2 vs M3
CFI	1.000	1.000	1.000
TLI	1.079	1.055	1.020
--2LL	26.454	24.537	22.170
H0	26.218	24.705	22.213
SCFMLR	1.049	1.135	1.154
		M1 vs M2	*M2 vs M3*
cd		0.878	1.079
TRd		3.446	4.619
pvalue		0.178	0.032

NONSTIMULUS ACTIVITY

Parameters	6.000	4.000	3.000
RMSEA	0.066	0.067	0.067
CFI	0.576	0.520	0.506
TLI	0.703	0.694	0.699
--2LL	-2021.322	-2023.341	-2024.091
H0	-2021.440	-2023.332	-2024.093
SCFMLR	0.690	0.950	0.936
		M1 vs M2	*M2 vs M3*
cd		0.171	0.993
TRd		22.194	1.533
pvalue		0.000	0.216

23

Table 3. Parameter estimates of the best fitting growth model for each neurophysiological measure include mean estimate values (Est.), standard errors of the mean (SE), and p-values. Intercepts reflect estimates at age 3. For example, at age 3, the average latency value was 90.744 ms (+/- 0.3 ms) and decreased by -0.018 (+/- 0.005) per year of age. Please note, for both spectral coding and nonstimulus activity, results were analyzed in nanovolts instead of microvolts due to small variance estimates. Structural equation models, like all methods derived from Generalized Linear Model (GLM), are invariant to linear transformations; results are the same if analyzed in nanovolts, microvolts, etc. To convert the estimate values from nanovolts to microvolts, means are divided by 1×10^3 and variance terms are divided by 1×10^6.

	NEURAL TIMING	SPECTRAL CODING	RESPONSE STABILITY	NONSTIMULUS ACTIVITY
Model with best fit	M2	M1	M2	---
Sample size	175	175	173	175
Slope with Intercept	---	Est. = - 4.167 SE = 1.796 p-value = 0.02	---	Est. = - 36.932 SE = 12.058 p-value = 0.002
Intercept Mean	Est. = 90.744 SE = 0.03 p-value < 0.001	Est. = 17.463 SE = 0.848 p-value < 0.001	Est. = 0.597 SE = 0.045 p-value < 0.001	Est. = 100.942 SE = 3.803 p-value < 0.001
Slope Mean	Est. = - 0.018 SE = 0.005 p-value < 0.001	Est. = 0.318 SE = 0.151 p-value = 0.035	Est. = 0.016 SE = 0.008 p-value = 0.043	Est. = - 0.765 SE = 0.648 p-value = 0.238
Intercept Variance	Est. = 0.037 SE = 0.005 p-value < 0.001	Est. = 44.414 SE = 11.633 p-value < 0.001	Est. = 0.025 SE = 0.004 p-value < 0.001	Est. = 441.382 SE = 118.489 p-value < 0.001
Slope Variance	---	Est. = 0.691 SE = 0.308 p-value = 0.025	---	Est. = 3.351 SE = 1.229 p-value = 0.006
Residual Variance	Est. = 0.014 SE = 0.002 p-value < 0.001	Est. = 8.392 SE = 1.026 p-value < 0.001	Est. = 0.035 SE = 0.004 p-value < 0.001	Est. = 328.371 SE = 25.262 p-value < 0.001

Table 4. Pearson's correlations between neural measures, controlling for age. Provided are the Pearson correlation coefficients (r) and p-values (p) for each correlation.

|	Neural timing	Spectral coding	Response stability	Nonstimulus activity
Neural timing	---	---	---	---
Spectral coding	r = - 0.400 p < 0.001	---	---	---
Response stability	r = - 0.217 p < 0.001	r = 0.656 p < 0.001	---	---
Nonstimulus activity	r = - 0.086 p = 0.75	r = 0.158 p = 0.001	r = - 0.446 p = 0.001	---

Figure 1. The speech syllable [da] (A), a 170 milisecond (ms) consonant-vowel stimulus, was presented to the right ear. The frequency following response (FFR) to the [da] (B) reflects various aspects of sound processing, including neural timing (C), spectral coding (D), response stability (E), and nonstimulus activity (F).

Figure 2. Path diagram reflecting the structure of the growth curve model used to estimate auditory neurophysiological development over time. Four series of growth curve models were run; each corresponded to one neural measure of sound processing: neural timing, spectral coding, response stability, and nonstimulus activity. Included in the models were estimates of intercept (I) and slope (S) means and variances. Constraining intercept values to 1 allowed us to determine the neural measure's initial status (i.e. at age 3). Constraining slope values to age of test (Age 3, Age 4, etc.) allowed us to determine linear changes in growth over time.

Figure 3. Neural timing becomes faster over time within childhood. Plotted are the frequency following responses (FFRs) to speech in the time domain for all children at age 3 (darkest blue, top left) through age 8 (lightest grey, bottom left). Waveforms for the six age points are overlaid on the right.

Figure 4. Spectral coding becomes stronger over time in childhood. Top panel: frequency following responses (FFRs) to speech are plotted in the frequency domain for all children at age 3 (darkest blue) through age 8 (lightest grey). Bottom panel: spectral amplitudes at each frequency. Please note, the bottom panel's x-axes are scaled with 40 hertz (Hz) bins centered around each frequency/harmonic of interest [e.g., fundamental frequency (F0) x-axis: 80 Hz to 120 Hz]. Y-axis values are scaled to best illustrate effects of age.

Figure 5. Response stability improves over time in childhood. Plotted are two sub-averages (rep 1 and rep 2; 2000 sweeps each) of frequency following responses (FFRs) to speech. Response stability is calculated by taking the Pearson product-moment correlation coefficient (r) of these two sub-averages.
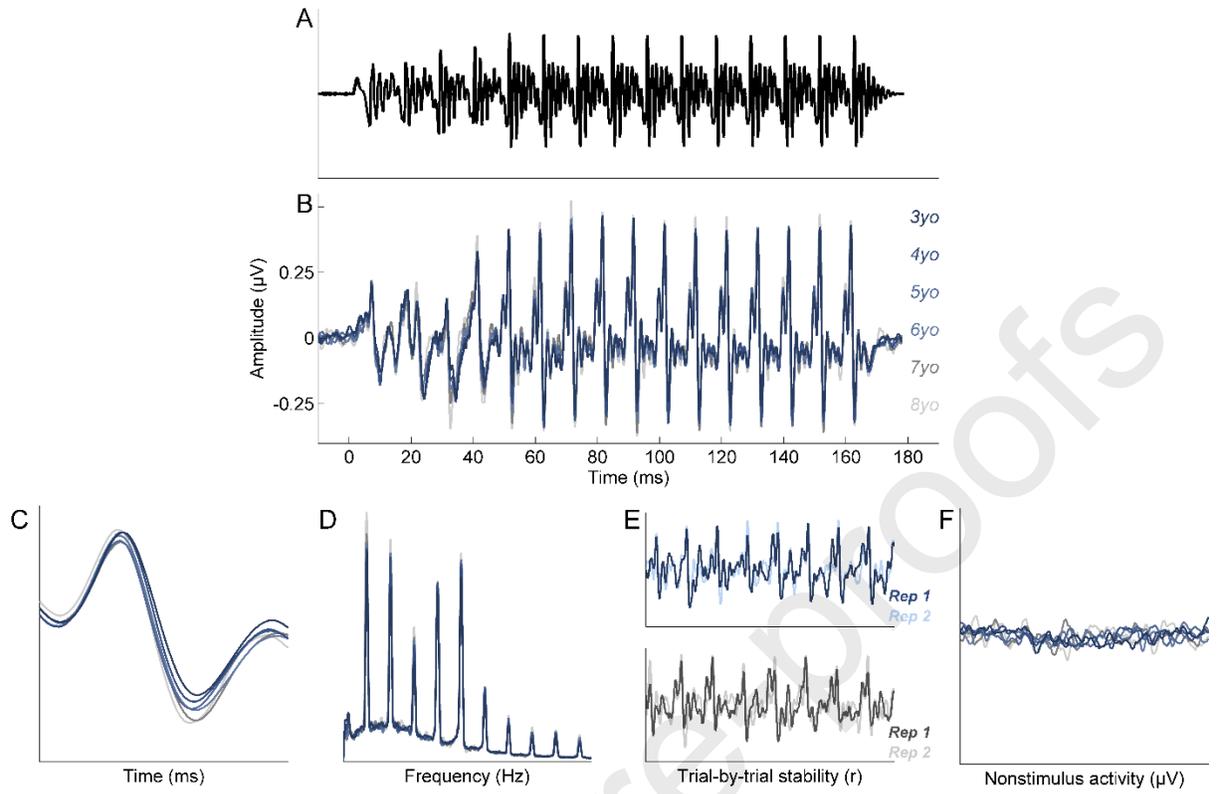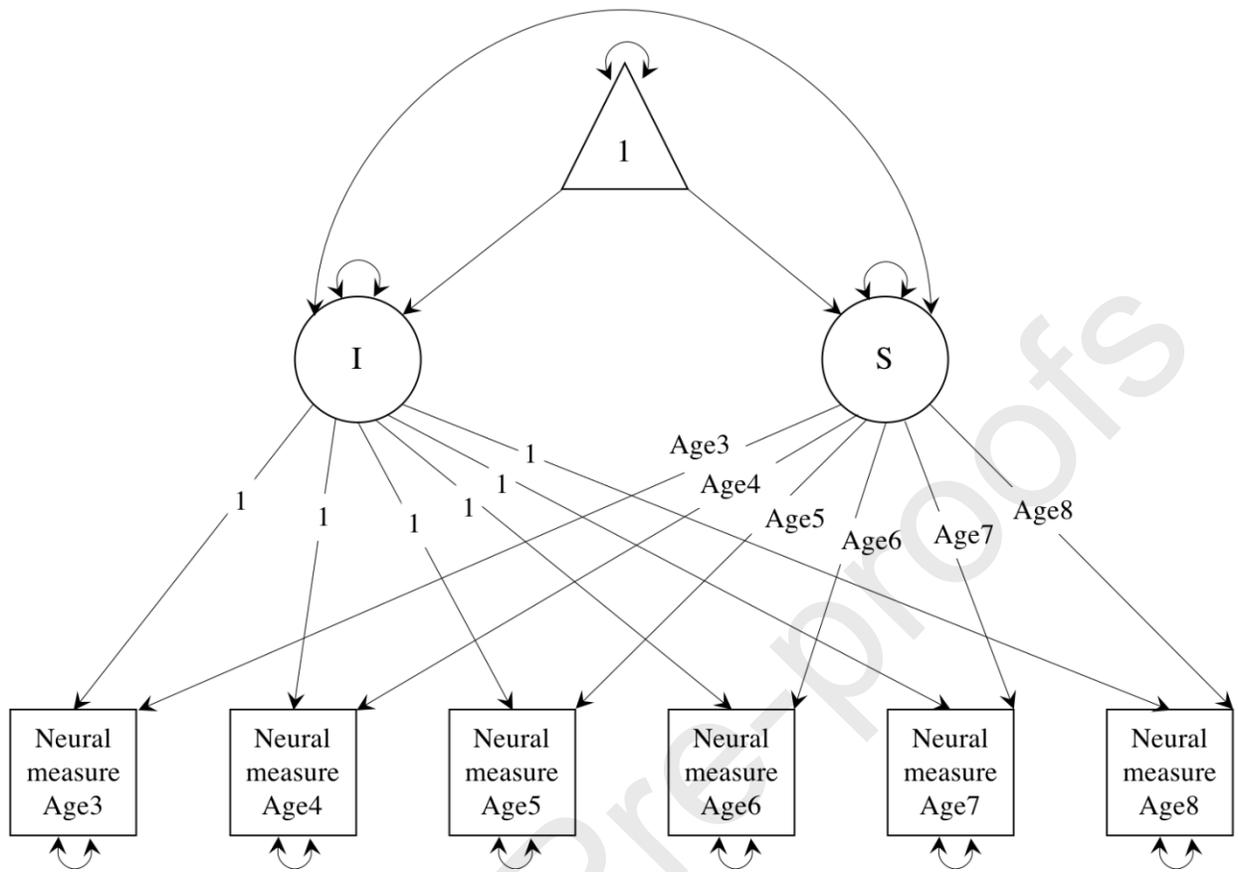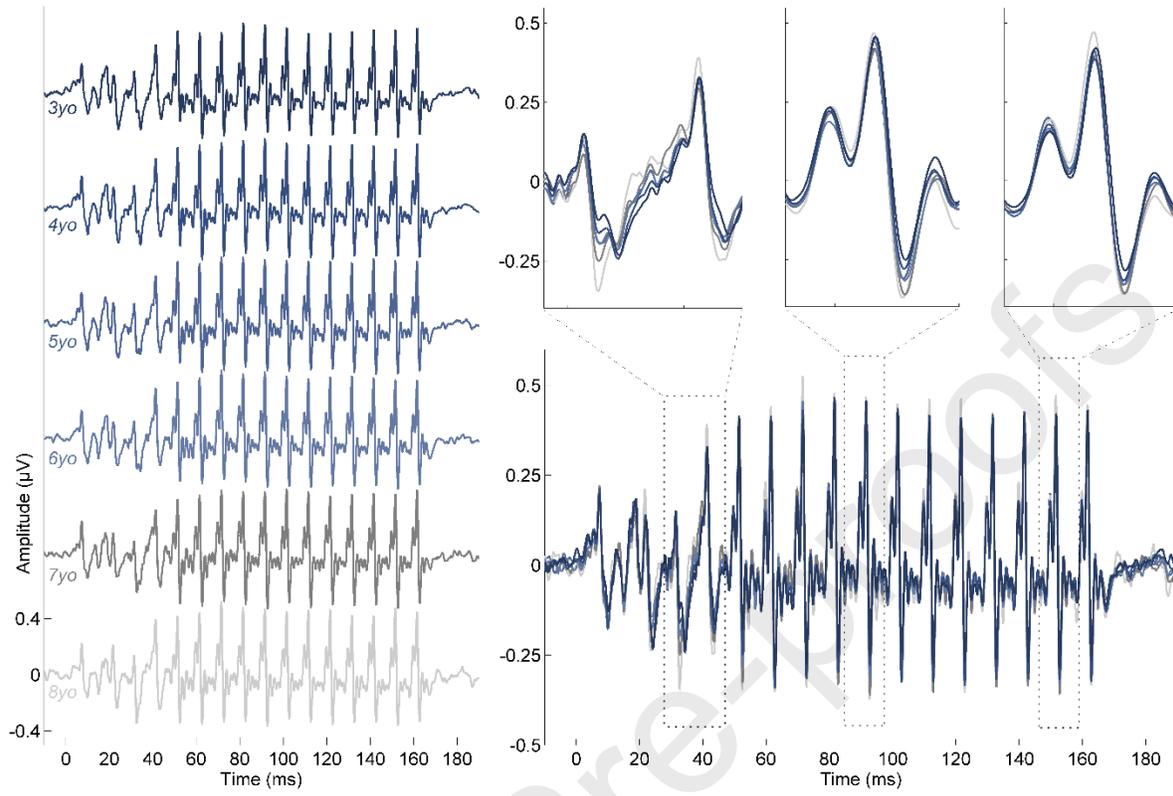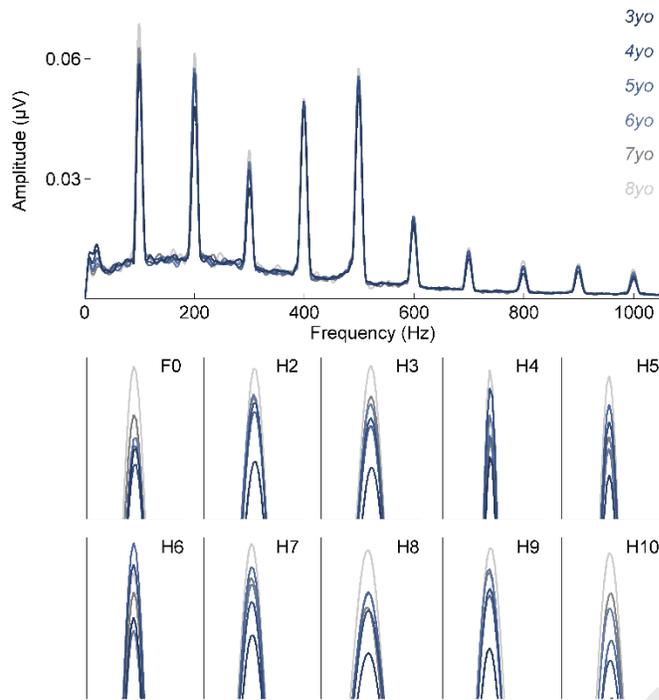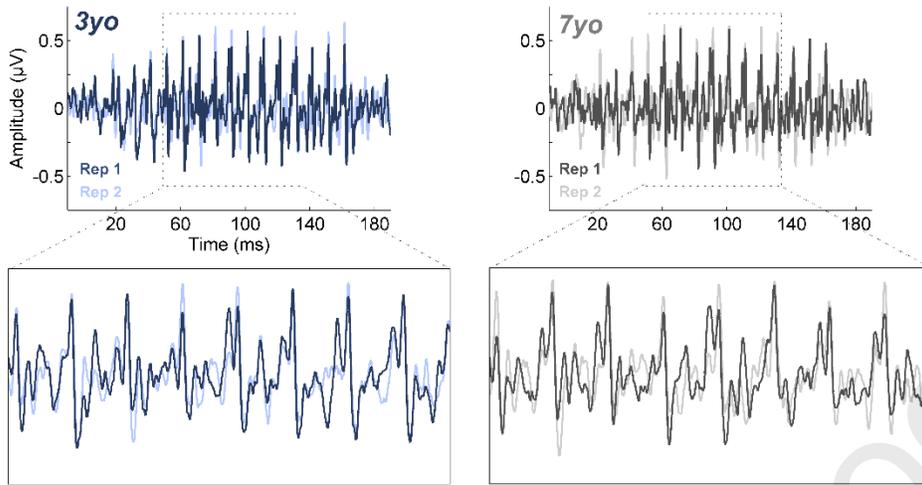
Appendix

Table A.1. *Parameter estimates for all growth curve models at each neurophysiological measure.*
In the main text (Table 3), only the best fitting growth curve model is reported for each neurophysiological measure. Here, we provide all growth curve model parameter estimates including mean estimate values (Est.), standard errors of the mean (SE), and p-values. Unshaded cells reflect estimates not reported in main text, while shaded cells are included in the main text. M1, M2, M3 refer to model 1, 2, and 3, respectively.

	M1	M2	M3	M1	M2	M3
	NEURAL TIMING			**SPECTRAL CODING**		
Sample size	175	175	175	175	175	175
Slope with Intercept	Est. = -0.002 SE = 0.004 p-value = 0.61	---	---	Est. = -4.167 SE = 1.796 p-value = 0.02	---	---
Intercept Mean	Est. = 90.748 SE = 0.033 p-value < 0.001	Est. = 90.744 SE = 0.03 p-value < 0.001	Est. = 90.656 SE = 0.016 p-value < 0.001	Est. = 17.463 SE = 0.848 p-value < 0.001	Est. = 17.55 SE = 0.87 p-value < 0.001	Est. = 19.006 SE = 0.38 p-value < 0.001
Slope Mean	Est. = -0.019 SE = 0.005 p-value < 0.001	Est. = -0.018 SE = 0.005 p-value < 0.001	---	Est. = 0.318 SE = 0.151 p-value = 0.035	Est. = 0.293 SE = 0.156 p-value = 0.06	---
Intercept Variance	Est. = 0.056 SE = 0.029 p-value < 0.001	Est. = 0.037 SE = 0.005 p-value < 0.001	Est. = 0.038 SE = 0.006 p-value < 0.001	Est. = 44.414 SE = 11.633 p-value < 0.001	Est. = 19.777 SE = 2.647 p-value < 0.001	Est. = 20.119 SE = 2.624 p-value < 0.001
Slope Variance	Est. = 0 SE = 0.001 p-value = 0.795	---	---	Est. = 0.691 SE = 0.308 p-value = 0.025	---	---
Residual Variance	Est. = 0.014 SE = 0.002 p-value < 0.001	Est. = 0.014 SE = 0.002 p-value < 0.001	Est. = 0.014 SE = 0.002 p-value < 0.001	Est. = 8.392 SE = 1.026 p-value < 0.001	Est. = 9.793 SE = 1.11 p-value < 0.001	Est. = 9.886 SE = 1.163 p-value < 0.001
	RESPONSE STABILITY			**NONSTIMULUS ACTIVITY**		
Sample size	173	173	173	175	175	175
Slope with Intercept	Est. = -0.009 SE = 0.006 p-value = 0.139	---	---	Est. = -36.932 SE = 12.058 p-value = 0.002	---	---
Intercept Mean	Est. = 0.591 SE = 0.045 p-value < 0.001	Est. = 0.597 SE = 0.045 p-value < 0.001	Est. = 0.679 SE = 0.016 p-value < 0.001	Est. = 100.942 SE = 3.803 p-value < 0.001	Est. = 101.225 SE = 3.866 p-value < 0.001	Est. = 96.973 SE = 1.321 p-value < 0.001
Slope Mean	Est. = 0.018 SE = 0.008 p-value = 0.031	Est. = 0.016 SE = 0.008 p-value = 0.043	---	Est. = -0.765 SE = 0.648 p-value = 0.238	Est. = -0.828 SE = 0.664 p-value = 0.212	---
Intercept Variance	Est. = 0.078 SE = 0.034 p-value = 0.022	Est. = 0.025 SE = 0.004 p-value < 0.001	Est. = 0.026 SE = 0.004 p-value < 0.001	Est. = 441.382 SE = 118.489 p-value < 0.001	Est. = 141.147 SE = 35.1 p-value < 0.001	Est. = 142.279 SE = 35.358 p-value < 0.001
Slope Variance	Est. = 0.002 SE = 0.001 p-value = 0.166	Est. = 0.078 SE = 0.034 p-value = 0.022	---	Est. = 3.351 SE = 1.229 p-value = 0.006	---	---
Residual Variance	Est. = 0.032 SE = 0.004 p-value < 0.001	Est. = 0.035 SE = 0.004 p-value < 0.001	Est. = 0.035 SE = 0.004 p-value < 0.001	Est. = 328.371 SE = 25.262 p-value < 0.001	Est. = 340.062 SE = 27.357 p-value < 0.001	Est. = 340.584 SE = 27.305 p-value < 0.001

Figure A.1. Effects of sex on neurophysiological development
Growth curve analyses were also performed including sex as a time-invariant factor in the models (Figure A.1) and are reported here due to the fact that sex was not a significant predictor of growth for the four neurophysiological measures. This is evident in Table A.2, where the inclusion of sex as a predictor had little to no impact on overall model fit, and in Table A.3, where Sex was not a significant predictor of slope (Sex on S, $p > 0.05$ for M1 and M2). Although we do not see an effect of sex on development, this does not necessarily indicate males and females mature at identical rates; rather, we do not have enough evidence to support the notion that there are sex differences across development.

We do see, however, a significant effect of intercept variance for all models, suggesting there are overall sex differences in neurophysiological processing of sound in this age range (Table A.3, mean sex $p < 0.05$ for all neural measures). In general, females had frequency following responses (FFRs) that were faster (neural timing), more robust spectrally (spectral coding), more consistent (response stability), and less noisy (nonstimulus activity) relative to their male peers.
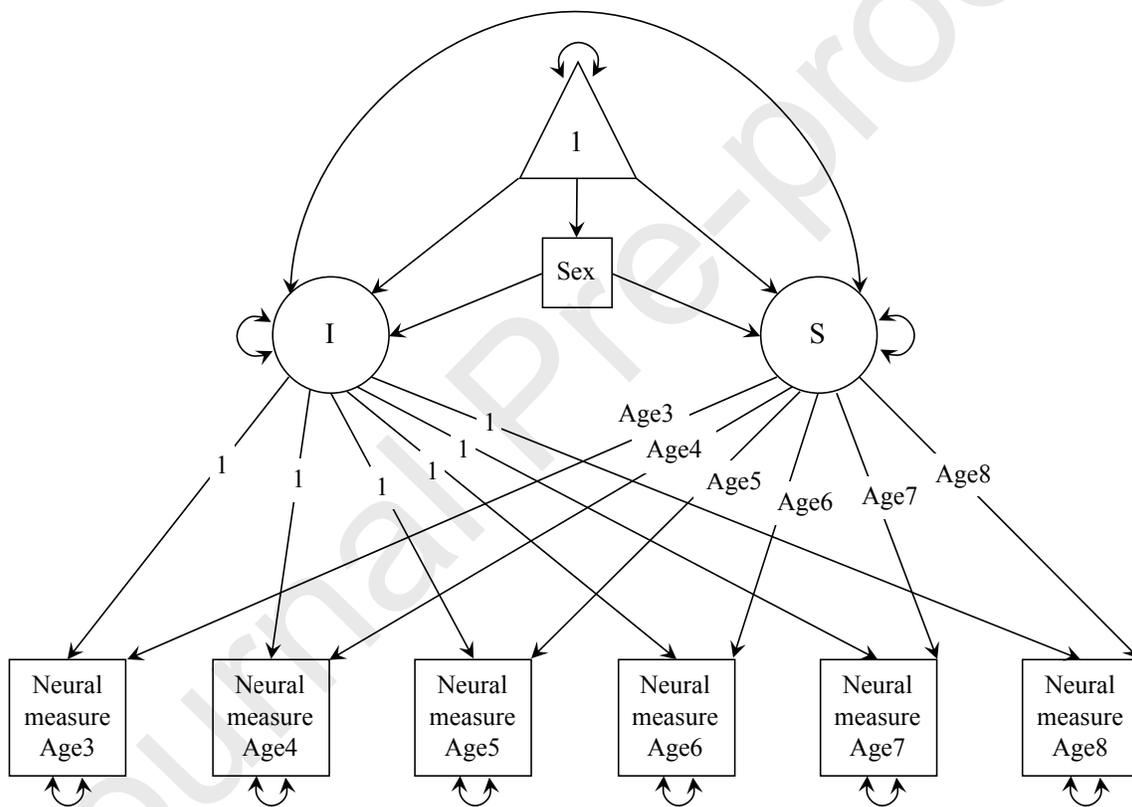
Figure A.1 Path diagram reflecting the structure of the growth curve model used to estimate auditory neurophysiological development over time. In a follow up series of growth curve models, sex was included as a time-invariant factor to investigate sex differences in auditory neurophysiological development. Four series of growth curve models were run, with each corresponding to one neural measure of sound processing: neural timing, spectral coding, response stability, and nonstimulus activity. Included in the models were estimates of intercept (I) and slope (S) means and variances. Constraining intercept values to 1 allowed us to determine the neural measure's initial status (i.e. at age 3). Constraining slope values to age of test (Age 3, Age 4, etc.) allowed us to determine linear changes in growth over time.

Table A.2. Fit statistics for growth curve models including sex as a time-invariant factor. Measures of model fit included Root Mean Square Error of Approximation (RMSEA), Comparative Fit Index (CFI), and Tucker-Lewis index (TLI). Unshaded cells represent model comparisons from the multi-level structural equation model where the basis coefficients for estimating change over time corresponded to age at test (e.g., 3.15 years, 3.87, etc.). The multi-level structural equation model comparisons (M1 vs M2 and M2 vs M3) were performed using Chi-Square Difference Tests based on log likelihood values (H0) and scaling correction factors obtained within the MLR estimator (SCFMLR), and were computed based on the difference test scaling correction (cd) and the Chi-Square difference test (TRd) values.

	M1	M2	M3		M1	M2	M3
NEURAL TIMING				***SPECTRAL CODING***			
Parameters	10	8	7	Parameters	10	8	7
RMSEA	0.1	0.1	0.111	RMSEA	0.037	0.049	0.047
CFI	0.851	0.839	0.793	CFI	0.97	0.943	0.946
TLI	0.876	0.876	0.847	TLI	0.975	0.956	0.96
--2LL	37.469	34.872	27.825	--2LL	-1410.965	-1414.621	-1414.783
H0	37.01	32.885	27.407	H0	-1410.888	-1414.199	-1414.369
SCFMLR	1.092	1.038	1.07	SCFMLR	0.933	1.045	1.011
		M1 vs M2	*M2 vs M3*			*M1 vs M2*	*M2 vs M3*
cd		1.309	0.814	cd		0.483	1.284
TRd		6.304	13.464	TRd		13.719	0.265
p-value		0.043	0.000	p-value		0.001	0.607
RESPONSE STABILITY				***NONSTIMULUS ACTIVITY***			
Parameters	10	8	7	Parameters	10	8	7
RMSEA	0	0	0	RMSEA	0.063	0.064	0.066
CFI	1	1	1	CFI	0.562	0.496	0.444
TLI	1.094	1.071	1.052	TLI	0.635	0.613	0.588
--2LL	-98.722	-100.462	-101.746	--2LL	-2146.737	-2148.972	-2150.457
H0	-98.924	-100.278	-101.481	H0	-2147.266	-2149.216	-2150.422
SCFMLR	0.945	0.963	0.942	SCFMLR	0.689	0.853	0.854
		M1 vs M2	*M2 vs M3*			*M1 vs M2*	*M2 vs M3*
cd		0.873	1.11	cd		0.032	0.847
TRd		3.101	2.168	TRd		120.37	2.849
p-value		0.212	0.141	p-value		0.000	0.091

Table A.3. Parameter estimates for growth curve models including sex as a time-invariant factor. Included are mean estimate values (Est.), standard errors of the mean (SE), and p-values. Included in the models were estimates of intercept (I) and slope (S) means and variances.

NEURAL TIMING	M1	M2	M3
Sample size	175	175	175
i on sex	Est. = 0.042 SE = 0.058 p-value = 0.469	Est. = 0.065 SE = 0.059 p-value = 0.276	Est. = 0.18 SE = 0.052 p-value < 0.001
s on sex	Est. = 0.016 SE = 0.009 p-value = 0.089	Est. = 0.01 SE = 0.009 p-value = 0.278	Est. = -0.014 SE = 0.007 p-value = 0.032
s with i	Est. = -0.003 SE = 0.005 p-value = 0.483	---	---
mean sex	Est. = 0.537 SE = 0.038 p-value < 0.001	Est. = 0.537 SE = 0.038 p-value < 0.001	Est. = 0.537 SE = 0.038 p-value < 0.001
intercept i	Est. = 90.731 SE = 0.043 p-value < 0.001	Est. = 90.712 SE = 0.038 p-value < 0.001	Est. = 90.596 SE = 0.023 p-value < 0.001
intercept s	Est. = -0.029 SE = 0.008 p-value < 0.001	Est. = -0.025 SE = 0.007 p-value < 0.001	---
variance s	Est. = 0.249 SE = 0.003 p-value < 0.001	Est. = 0.249 SE = 0.003 p-value < 0.001	Est. = 0.249 SE = 0.003 p-value < 0.001
residual variance neural timing	Est. = 0.014 SE = 0.002 p-value < 0.001	Est. = 0.014 SE = 0.002 p-value < 0.001	Est. = 0.014 SE = 0.002 p-value < 0.001
residual variance i	Est. = 0.062 SE = 0.032 p-value = 0.052	Est. = 0.033 SE = 0.005 p-value < 0.001	Est. = 0.034 SE = 0.005 p-value < 0.001
residual variance s	Est. = 0 SE = 0.001 p-value = 0.772	---	---

SPECTRAL CODING	M1	M2	M3
Sample size	175	175	175
i on sex	Est. = -2.979 SE = 1.643 p-value = 0.07	Est. = -3.012 SE = 1.702 p-value = 0.077	Est. = -3.565 SE = 1.246 p-value = 0.004
s on sex	Est. = 0.33 SE = 0.289 p-value = 0.253	Est. = 0.333 SE = 0.303 p-value = 0.271	Est. = 0.448 SE = 0.206 p-value = 0.03
s with i	Est. = -3.867 SE = 1.733 p-value = 0.026	---	---
mean sex	Est. = 0.537 SE = 0.038 p-value < 0.001	Est. = 0.537 SE = 0.038 p-value < 0.001	Est. = 0.537 SE = 0.038 p-value < 0.001
intercept i	Est. = 19.044 SE = 1.264 p-value < 0.001	Est. = 19.151 SE = 1.318 p-value < 0.001	Est. = 19.705 SE = 0.579 p-value < 0.001
intercept s	Est. = 0.143 SE = 0.215 p-value = 0.508	Est. = 0.115 SE = 0.228 p-value = 0.613	---
variance s	Est. = 0.249 SE = 0.003 p-value < 0.001	Est. = 0.249 SE = 0.003 p-value < 0.001	Est. = 0.249 SE = 0.003 p-value < 0.001
residual variance spectral coding	Est. = 8.41 SE = 1.03 p-value < 0.001	Est. = 9.756 SE = 1.12 p-value < 0.001	Est. = 9.757 SE = 1.129 p-value < 0.001
residual variance i	Est. = 41.71 SE = 10.909 p-value < 0.001	Est. = 19.244 SE = 2.536 p-value < 0.001	Est. = 19.287 SE = 2.513 p-value < 0.001
residual variance s	Est. = 0.66 SE = 0.306 p-value = 0.031	---	---

RESPONSE STABILITY	M1	M2	M3
Sample size	175	175	175
i on sex	Est. = -0.065 SE = 0.089 p-value = 0.468	Est. = -0.059 SE = 0.087 p-value = 0.499	Est. = -0.141 SE = 0.065 p-value = 0.031
s on sex	Est. = 0.002 SE = 0.016 p-value = 0.891	Est. = 0.001 SE = 0.016 p-value = 0.974	Est. = 0.017 SE = 0.011 p-value = 0.113
s with i	Est. = -0.008 SE = 0.006 p-value = 0.166	---	---
mean sex	Est. = 0.537 SE = 0.038 p-value < 0.001	Est. = 0.537 SE = 0.038 p-value < 0.001	Est. = 0.537 SE = 0.038 p-value < 0.001
intercept i	Est. = 0.624 SE = 0.065 p-value < 0.001	Est. = 0.626 SE = 0.064 p-value < 0.001	Est. = 0.708 SE = 0.024 p-value < 0.001
intercept s	Est. = 0.017 SE = 0.012 p-value = 0.151	Est. = 0.017 SE = 0.012 p-value = 0.152	---
variance s	Est. = 0.249 SE = 0.003 p-value < 0.001	Est. = 0.249 SE = 0.003 p-value < 0.001	Est. = 0.249 SE = 0.003 p-value < 0.001
residual variance response stability	Est. = 0.032 SE = 0.004 p-value < 0.001	Est. = 0.035 SE = 0.004 p-value < 0.001	Est. = 0.035 SE = 0.004 p-value < 0.001
residual variance i	Est. = 0.074 SE = 0.033 p-value = 0.027	Est. = 0.024 SE = 0.004 p-value < 0.001	Est. = 0.024 SE = 0.004 p-value < 0.001
residual variance s	Est. = 0.001 SE = 0.001 p-value = 0.204	---	---

NONSTIMULUS ACTIVITY	M1	M2	M3
Sample size	175	175	175
i on sex	Est. = -4.523 SE = 7.513 p-value = 0.547	Est. = -4.962 SE = 7.62 p-value = 0.515	Est. = 2.719 SE = 5.847 p-value = 0.642
s on sex	Est. = 1.177 SE = 1.271 p-value = 0.354	Est. = 1.284 SE = 1.298 p-value = 0.322	Est. = -0.245 SE = 0.958 p-value = 0.798
s with i	Est. = -34.254 SE = 8.207 p-value < 0.001	---	---
mean sex	Est. = 0.537 SE = 0.038 p-value < 0.001	Est. = 0.537 SE = 0.038 p-value < 0.001	Est. = 0.537 SE = 0.038 p-value = 0.000
intercept i	Est. = 103.394 SE = 5.232 p-value < 0.001	Est. = 103.974 SE = 5.297 p-value < 0.001	Est. = 96.167 SE = 2.012 p-value < 0.001
intercept s	Est. = -1.412 SE = 0.867 p-value = 0.104	Est. = -1.548 SE = 0.877 p-value = 0.077	---
variance s	Est. = 0.249 SE = 0.003 p-value < 0.001	Est. = 0.249 SE = 0.003 p-value < 0.001	Est. = 0.249 SE = 0.003 p-value < 0.001
residual variance nonstimulus activity	Est. = 327.831 SE = 25.229 p-value < 0.001	Est. = 338.553 SE = 27.633 p-value < 0.001	Est. = 341.98 SE = 27.536 p-value = 0.000
residual variance i	Est. = 428.576 SE = 101.432 p-value < 0.001	Est. = 141.683 SE = 36.609 p-value < 0.001	Est. = 139.308 SE = 35.588 p-value < 0.001
residual variance s	Est. = 2.826 SE = 0.668 p-value < 0.001	---	---