Research Article

Psychological Science
1–13

DOI: 10.1177/0956797620933237
www.psychologicalscience.org/PS

Preverbal Infants Discover Statistical Word Patterns at Similar Rates as Adults: Evidence From Neural Entrainment

TC

Dawoon Choi[1], Laura J. Batterink[2,3], Alexis K. Black[4], Ken A. Paller[5], and Janet F. Werker[1]

[1]Department of Psychology, University of British Columbia; [2]Department of Psychology, Western University; [3]The Brain and Mind Institute, Western University; [4]School of Audiology and Speech Sciences, University of British Columbia; and [5]Department of Psychology, Northwestern University

Abstract

The discovery of words in continuous speech is one of the first challenges faced by infants during language acquisition. This process is partially facilitated by statistical learning, the ability to discover and encode relevant patterns in the environment. Here, we used an electroencephalogram (EEG) index of neural entrainment to track 6-month-olds' (N = 25) segmentation of words from continuous speech. Infants' neural entrainment to embedded words increased logarithmically over the learning period, consistent with a perceptual shift from isolated syllables to wordlike units. Moreover, infants' neural entrainment during learning predicted postlearning behavioral measures of word discrimination (n = 18). Finally, the logarithmic increase in entrainment to words was comparable in infants and adults, suggesting that infants and adults follow similar learning trajectories when tracking probability information among speech sounds. Statistical-learning effects in infants and adults may reflect overlapping neural mechanisms, which emerge early in life and are maintained throughout the life span.

Keywords

speech perception, language development, learning, infant development, open data, open materials, preregistered

Received 10/30/19; Revision accepted 5/5/20

One of the first challenges faced by infants during language acquisition is the discovery of word boundaries in continuous speech. This process, known as *word segmentation*, is not trivial, as the speech signal does not contain reliable phonetic cues to word boundaries (e.g., Klatt, 1980). Despite the computational challenges of word segmentation, infants have already learned some word forms of their native language by 6 months of age (Bergelson & Swingley, 2012). *Statistical learning*, the process of becoming sensitive to statistical regularities in the environment, has been proposed as a learning mechanism that supports the initial process of word segmentation. In a given language, the probability of two syllables co-occurring is higher within than across word boundaries (Swingley, 2005). Tracking this probability information may be one mechanism by which infants segment words from natural speech (Saffran, Aslin, & Newport, 1996; Saffran & Kirkham, 2018).

Saffran et al. (1996) provided the first evidence that infants can extract wordlike units by tracking the statistical patterns between speech sounds. Saffran et al. used an artificial language that was created by concatenating four trisyllabic nonsense words (e.g., *babupu*, *padoti*) in a pseudorandom sequence without any acoustic cues to word boundaries. After 2 min of exposure to this language, infants discriminated words from

Corresponding Authors:
Dawoon Choi, University of British Columbia, Department of Psychology, 2136 West Mall, Vancouver, British Columbia, Canada, V6T 1Z4
E-mail: dchoi@psych.ubc.ca

Laura J. Batterink, The Brain and Mind Institute, Western University, Western Interdisciplinary Research Building, Room 6124, London, Ontario, Canada N6A 3K7
E-mail: lbatter@uwo.ca

nonword foils (trisyllabic sequences composed of recombined syllables), as reflected by longer listening times to nonwords. Later work has demonstrated that infants are sensitive not only to probabilistic sequences of syllables but also to nonlinguistic auditory and visual stimuli (for a review, see Saffran & Kirkham, 2018). There is also evidence that statistical word segmentation facilitates other aspects of language acquisition, such as vocabulary learning (Graf Estes, Evans, Alibali, & Saffran, 2007). Neuroimaging studies of newborns have provided evidence that a statistical-learning ability for speech segmentation is already present at birth (Fló et al., 2019; Teinonen, Fellman, Näätänen, Alku, & Huotilainen, 2009). Taken together, these results suggest that infants can extract embedded words from continuous speech by tracking the statistical probabilities between syllables.

To date, most evidence for statistical learning in infants comes from postlearning behavioral tests, in which infants' responses to both previously encountered words and novel foil items are measured after familiarization. Although this approach has yielded insights into infants' statistical-learning abilities, it has several limitations. First, the infant's response to the segmented words reflects the outcome of learning but not statistical learning itself. Thus, test performance may be influenced by extrinsic factors, such as general interest in the testing procedure, motor abilities, and memory retrieval. Second, postlearning tests cannot reveal the time course or the dynamics of learning. Previous research has shown that infants can learn syllabic chunks from speech streams after variable lengths of familiarization, but little is known about when an infant becomes sensitive to the underlying statistical structure. Finally, behavioral tests using measures of looking-time preference are not suitable beyond infancy, precluding meaningful comparisons of statistical learning across the life span. Consequently, the different methodologies used for infants and for older children and adults impede our understanding of how the mechanisms and dynamics of statistical learning change with maturation and experience.

In the present study, we aimed to overcome these limitations by using a neural index of statistical learning to track learners' sensitivity to hidden embedded words during familiarization using a measure of neural entrainment based on electroencephalograms (EEGs; e.g., Batterink & Paller, 2017, 2019; Buiatti, Peña, & Dehaene-Lambertz, 2009; Kabdebon, Pena, Buiatti, & Dehaene-Lambertz, 2015). *Neural entrainment* refers to the synchrony between neural oscillations and external stimuli that emerges during the processing of rhythmic or predictable visual and auditory stimuli, including speech. For example, electrophysiological measures of speech processing reveal peaks in time-locked cortical activity at different linguistic units, including single syllables, words, and phrases (Ding, Melloni, Zhang, Tian, & Poeppel, 2016). Importantly, neural entrainment does not merely reflect acoustic features of the signal but is sensitive to higher-level cognitive processes, such as listeners' knowledge and subjective perception as guided by syntactic rules (Ding et al., 2016) or imagined rhythms (Nozaradan, Peretz, Missal, & Mouraux, 2011). In the context of statistical learning, the central assumption is that as learners gradually discover words in continuous speech, neuronal firing should progressively entrain more strongly to the word frequency relative to the low-level syllable frequency, representing a shift in perception from syllables to words (Buiatti et al., 2009).

Consistent with the hypothesis that there is a relationship between neural entrainment and learning of the probabilistic relationships between syllables, findings of recent EEG studies of statistical learning in adults have confirmed that neural entrainment can reveal emergent word-sized units and that such entrainment predicts word recognition on a subsequent behavioral test (Batterink & Paller, 2017, 2019). These results indicate that an EEG-based measure of neural entrainment can track statistical learning as it occurs. This methodology was also used to investigate the statistical learning of nonadjacent structures (Buiatti et al., 2009; Kabdebon et al., 2015), although in these latter studies, the insertion of brief pauses between words provided a low-level cue to segmentation.

In the present study, we applied the EEG frequency-tagging approach to study statistical learning in preverbal 6-month-old infants. We exposed infants to an artificial language similar to one developed by Saffran et al. (1996), in which trisyllabic nonsense words were concatenated without acoustic cues to word boundaries. To examine the temporal dynamics of statistical learning, we measured infants' EEG during familiarization and used a sliding-time-window analysis to index the trajectory of neural entrainment at a fine-grained scale. We hypothesized that there is a relationship between statistical learning and synchronization of ongoing neural oscillations to the embedded word structure. We therefore predicted that although infants would entrain to the individual syllables from the beginning, over the course of familiarization, infants would show a relative increase in word-level entrainment, reflecting a progressive acquisition of word knowledge through statistical learning.

After familiarization, we used an established behavioral paradigm to assess infants' ability to discriminate previously encountered words from novel nonwords. We predicted that word-level entrainment would correlate with the magnitude of infants' preference for nonwords. Such a finding would be consistent with previous findings in adults (Batterink & Paller, 2017, 2019), in which neural entrainment observed during exposure predicted later behavioral measures of statistical learning. Finally, we compared these infant data with existing adult data that were previously collected under comparable conditions (Batterink & Paller, 2017). We investigated the learning trajectory in each group to test whether infants and adults show differences in the rate or efficiency of statistical learning.

Method

Participants

Twenty-five full-term 6-month-old monolingual English-learning infants (age: M = 175.4 days, SD = 10.4 days) from the greater Vancouver, Canada, area were included in the familiarization-phase analyses. An additional seven infants were tested but excluded from analysis because of behavioral issues (n = 5) and equipment or experimenter error (n = 2). Sample size was estimated on the basis of effect sizes from adult data (N = 45) reported in Batterink and Paller's (2017) study, which produced an estimated Cohen's d effect size of 0.56 and observed power of .98 for the difference in the word-learning index (WLI) between structured and random speech streams. (The WLI is a neural-entrainment measure of relative sensitivity to the trisyllabic word structure; see the EEG Recording and Analysis section for more details.) Given the high similarity in design between Batterink and Paller's work (which was based on an adult sample) and the present study, as well as the fact that neural indicators of speech perception have been reported to be highly similar between infants and adults (e.g., Dehaene-Lambertz & Gliga, 2004), we determined that Batterink and Paller's study would be the best one to date on which to base our power estimates. We estimated that a comparable effect and an estimated power of .80 would require a minimum sample size of 22 infants (G*Power 3.1; Faul, Erdfelder, Lang, & Buchner, 2007). We selected 24 infants as a target sample size because that paralleled the sample size used in the original Saffran et al. study (1996). One additional infant was recruited during the recruitment window, leaving a final sample of 25 infants in the EEG entrainment analysis. Eighteen of the 25 infants yielded sufficient test-phase data (see the Supplemental Material available online for details). The research was approved by the University of British Columbia Behavioural Research Ethics Board (Ethics Certificate H95-80023).

Stimuli

Familiarization phase. The stimuli consisted of 12 consonant-vowel syllables (duration: 185–228 ms) that followed the phonemic conventions of English. The syllables were arranged into four trisyllabic words. To control for potential idiosyncratic stimulus effects, we constructed two separate artificial languages from the same 12 syllables. Each infant was randomly assigned to listen to one of the two artificial languages (see Table S1 in the Supplemental Material).

The individual syllables were generated using an artificial speech synthesizer (built-in Mac text-to-speech application; voice: "Samantha") and recorded in Audacity (https://www.audacityteam.org/) at a sampling rate of 44100 Hz. During familiarization, infants heard a continuous auditory stream (minimum 120 s, maximum 400 s) generated by concatenating the four trisyllabic words in a pseudorandom order without pauses. Syllable-to-syllable onset was 300 ms (3.3 Hz); word-to-word onset was 900 ms (1.1 Hz). The boundaries between words were cued only by the transitional probabilities between the syllables (i.e., the probability that two syllables would co-occur based on the overall frequency of the first syllable; 1.0 within a word and 0.33 across words).

Test phase. In the test phase, infants heard two words from the artificial language and two nonword foils. For each trial, infants heard repetitions of the test item separated by 1,100 ms of silence. The syllable-to-syllable onset within each word was 300 ms, as in the familiarization phase.

Procedure

The experimental protocol consisted of two phases: an initial familiarization phase and a test phase. Throughout the experiment, the infant sat on his or her caregiver's lap in a sound-attenuated booth. Caregivers listened to masking music through a pair of headphones to prevent them from hearing the auditory stimuli.

Familiarization phase. Infants were exposed to the artificial language playing over speakers (total number of syllables presented: 602–1,317; M = 1,123; SD = 201). A computer monitor in front of the infant displayed a dynamic abstract shape to hold the infants' attention. We implemented a minimum 120-s exposure criterion because previous studies have shown that about 2 min of exposure is sufficient for infants to learn (Kabdebon et al., 2015; Saffran et al., 1996). We set the maximum exposure duration to 360 s (corresponding to 1,200 syllables) in order to maximize the amount of data for EEG analyses, although because of experimenter error, several infants received slightly more exposure (up to 400 s; see Fig. S1 in the Supplemental Material). If the infant began to vocalize or move excessively and severe EEG artifacts

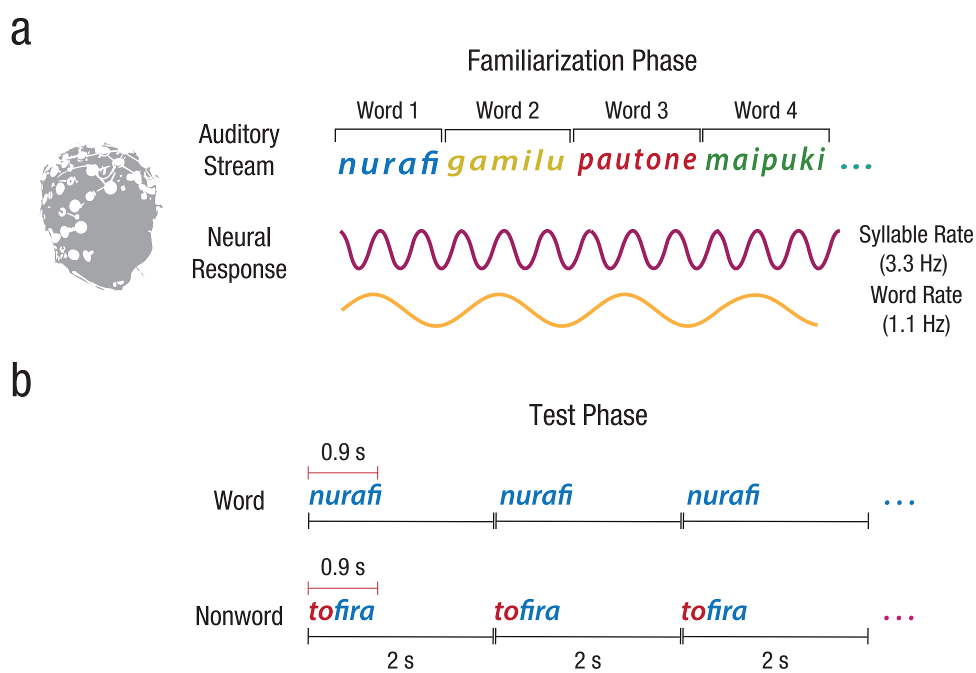

Fig. 1. Summary of the experimental protocol. The experiment consisted of a familiarization phase followed by a test phase. The familiarization phase (a) consisted of passive listening to an artificial language composed of repeating trisyllabic nonsense words. Neural entrainment to the auditory speech stream was measured via scalp electrodes throughout familiarization. In the test phase (b), each trial was composed of a single repeating test item (word or nonword), and infants controlled the duration of the trial by keeping their eyes focused on a rotating spiral. Infants' looking times to words and nonwords were measured. The different colors of the nonword syllables indicate the word in which the syllables were embedded during the familiarization phase.

were observed, or if the caregiver signaled to stop the speech stream, the experimenter paused the familiarization phase to check on the infant and the setup (pause: M = 33.2 s; SD = 19.0; n = 6). Familiarization resumed if the issue could be quickly addressed (n = 5); otherwise, the familiarization phase was terminated (n = 1). The infants progressed to the test phase following familiarization if they had completed a minimum of 2 min of continuous exposure and were calm enough to continue to the test phase, as judged by the caregiver.

Test phase. We used a central-fixation looking-time paradigm (Shi & Werker, 2001) to test infants' ability to discriminate words (trisyllabic structures embedded in the artificial language) and nonword foils (novel trisyllabic structures that infants had not encountered from the artificial language; see Table S1 in the Supplemental Material). Sequences of repeating test items (words and nonwords) were paired with a rotating spiral displayed on the screen. Infants controlled the duration of the repeating auditory stimuli through their looking behavior: A sustained shift in visual attention away from the screen terminated the trial. Infants' looking time to each trial represented our key measure of statistical learning. We predicted longer looking times to foil items, reflecting a novelty preference, as has been found in most previous statistical-learning research using similar paradigms (Graf Estes & Lew-Williams, 2015; Saffran et al., 1996).

A camera mounted above the computer monitor allowed the experimenter to observe the infant. A yellow ball preceded each trial to direct the infant's attention to the screen. When the infant looked at the screen, the experimenter, who was blind to the trial condition, initiated the trial presentation. On each trial, the infant heard a repetition of either a word or a nonword. The experimenter monitored and indicated on-line when the infant's gaze moved away from or returned to the screen. The trial ended when the infant looked away from the screen for at least 2 consecutive seconds or after a maximum trial duration of 30 s. Trials were presented pseudorandomly so that infants heard each of the two words and two nonwords every four trials, which constituted a block. Infants progressed through the test items until no longer compliant or until they completed a maximum of 16 trials. Infants who completed at least one test block were included. Figure 1 summarizes the experimental protocol.

EEG recording and analysis

EEG data were collected throughout the familiarization and test phases at a sampling rate of 1000 Hz with the

64-electrode geodesic sensor net (Net Amps 400 amplifier; Electrical Geodesics, Eugene, OR) referenced to the vertex (Cz). The maximal impedance was kept under 40 kΩ.

Quantification of neural entrainment across the familiarization phase

EEG neural-entrainment analyses were carried out using EEGLAB (Delorme & Makeig, 2004) and followed the same general procedure as in previous studies (Batterink & Paller, 2017, 2019). Some modifications were made to the artifact-correction method to accommodate the infant EEG data, which tend to have different artifact profiles compared with adult data (e.g., more motion-related artifacts). First, a 60-Hz notch filter and a band-pass filter from 0.5 to 20 Hz were applied to the raw data. Next, following a previous study of neural entrainment in infants (Cirelli, Spinelli, Nozaradan, & Trainor, 2016), we performed artifact correction using the artifact-block algorithm in MATLAB (The MathWorks, Natick, MA; Mourad, Reilly, de Bruin, Hasey, & MacCrimmon, 2007). The artifact-block algorithm improves the signal-to-noise ratio in infant data relative to conventional trial rejection and can be applied to both continuous and epoched EEGs (Fujioka, Mourad, He, & Trainor, 2011). As recommended by Fujioka et al. (2011), we applied the algorithm after removing the outer ring of channels from further analysis, and we set the threshold θ within the artifact-block algorithm at ±50 μV. The artifact-corrected data were then rereferenced to a common average, and data epochs of 9.0 s were extracted. Epochs were time locked to the onset of each word and corresponded to the duration of 10 trisyllabic words or 30 syllables (with no prestimulus interval); neighboring epochs overlapped for 9/10ths of their length. Visual inspection of the individual data confirmed that the artifact-block algorithm successfully corrected artifacts and reduced noise in the data.

We quantified neural entrainment by measuring intertrial coherence (ITC), a measure of event-related phase locking. ITC values range from 0, indicating purely non-phase-locked activity, to 1, indicating strictly phase-locked activity. A significant ITC indicated that the EEG activity in single trials was phase locked at a given time and frequency, rather than phase random with respect to the time-locking experimental event. ITC was computed using a continuous Morlet wavelet transformation from 0.555 to 5.0 Hz via the “newtimef” function of EEGLAB. Wavelet transformations were computed in 0.11-Hz steps, with two cycles at the lowest frequency (0.555 Hz) increasing by a scaling factor of 0.5 and reaching nine cycles at the highest frequency (5.0 Hz). This approach was selected to optimize the trade-off between temporal resolution at low frequencies and frequency resolution at high frequencies (Delorme & Makeig, 2004). As an initial step, ITC was computed across all epochs for each infant and plotted as a function of frequency (Fig. 2a; topographical voltage map represents ITC from all scalp electrodes retained after artifact-block correction). We selected a group of 23 frontocentral electrodes for all subsequent ITC computations.

Electrode selection was based on (a) visual inspection of the topographical distribution of the entrainment effect in our sample, (b) the general finding that auditory event-related potentials are mainly frontocentrally distributed, and (c) our previous finding with adults showing that entrainment effects are maximal at frontocentral sites (Batterink & Paller, 2017). Note that strong entrainment was also observed over posterior electrodes (Fig. 2a), despite the fact that auditory EEG signals generally show a frontocentral distribution; these channels were not included in ITC computations. This observed shift in the distribution of ITC relative to our original analysis in adults is likely due to the use of a common average reference instead of a mastoid reference, which results in the subtraction of higher amplitude EEG signals from the relatively weak signals over posterior regions. This use of an average reference was necessary because peripheral channels, which include the mastoids, are often noisy in infant data (Fujioka et al., 2011) and thus are not suitable references.

Neural entrainment time-course analysis (familiarization phase)

We predicted a relative increase in word-level neural entrainment over the course of exposure, reflecting a progression in sensitivity to the word structure in the speech stream. To test this hypothesis, we used a sliding-time-window analysis to examine the trajectory of learning at a fine-grained scale. We grouped every 10 consecutive epochs together into bundles (i.e., Epochs 1–10, Epochs 11–20); thus, each word was represented equally across bundles, except for nine words occurring at the beginning and nine words occurring at the end of the familiarization period. ITC was computed within each bundle of 10 consecutive epochs using the same “newtimef” function and parameters described previously. The WLI, a relative measure of sensitivity to the trisyllabic word structure, was then computed for each bundle by dividing ITC at the word frequency (1.1 Hz) with ITC at the syllable frequency (3.3 Hz; Batterink & Paller, 2017). Higher WLI values indicate greater neural entrainment toward the word frequency relative to the raw syllable frequency, which is indicative of statistical learning. The sliding-time-window analysis provides a

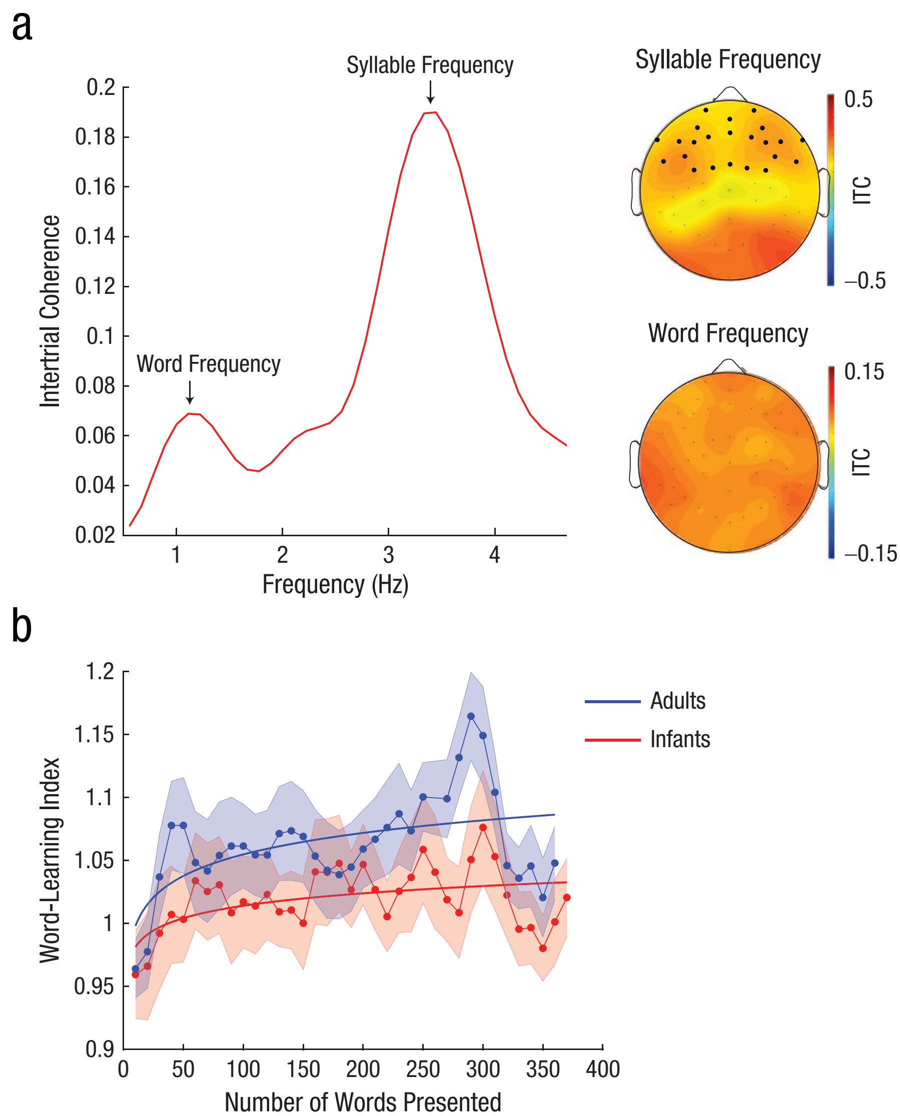

Fig. 2. Neural entrainment and statistical learning across the familiarization period. In (a), the graph shows intertrial coherence (ITC) as a function of frequency, and the topographical plots show the distribution of ITC values across the scalp. Peaks in neural entrainment at both the syllable and word frequencies are indicated in the graph. For the topographical plots, note that different scales are used for word and syllable frequencies. The dots on the upper scalp plot denote the approximate locations of the 23 centrofrontal electrodes used for all ITC analyses. Statistical learning, as assessed by the word-learning index (WLI), is shown in (b) as a function of the number of words presented, separately for infants and adults (adult data are from Batterink & Paller, 2017). Each line represents the best logarithmic fit for the group-averaged data. Shaded regions represent the standard error of each data point. Note that because ITC is highly influenced by trial number, ITC and the corresponding WLI values cannot be directly compared between (a) and (b) because ITC in (a) is computed across all epochs.

fine-grained (bundle-by-bundle) relative measure of word-level entrainment throughout the familiarization period. Because each bundle consists of only 10 epochs (resulting in relatively noisy WLI time-course data), data were smoothed using a moving-average filter with a span of three data points (i.e., each nth data point was averaged with data points $n - 1$ and $n + 1$). Because values for the first and final bundle for each participant cannot be smoothed, they were excluded from further analysis. The remaining smoothed WLI values were used in all subsequent statistical analyses.

WLI values for each bundle produced a learning curve over the familiarization period. At the group level, the learning curve was better characterized by a logarithmic function than a linear function (Fig. 2b), consistent with previous research on the time course of statistical learning in adults (Siegelman, Bogaerts, Kronenfeld, & Frost, 2018). Thus, to estimate statistical learning at the individual level, we modeled each infant's learning curve with a logarithmic function, $y = a + b \times \log(x)$, where y represents the WLI value and x represents the cumulative number of words each participant was exposed to during familiarization. Positive b values indicate that the WLI increases logarithmically as a function of exposure; larger b values reflect faster (steeper) learning rates. To examine our first hypothesis that infants would show a relative increase in word-level neural entrainment over the course of exposure, we used a one-sample t test to determine whether b was greater than 0, which would indicate an increase in WLI over the course of exposure to the speech stream. As a control, we assessed the distribution of randomly ordered bundles using a repeated random-sampling approach. For each participant, WLI values for each bundle were randomly shuffled and then smoothed using the same parameters and constraints as in the original analysis. A b value was calculated for each participant on the basis of these shuffled bundles, and an average b value across participants was then computed. This procedure was repeated a total of 1,000 times, producing a distribution of 1,000 surrogate b values for randomly ordered bundles. We then tested whether our observed b value (for the real data) fell within the upper 5% of this shuffled surrogate distribution.

Finally, for each infant, the logarithmic function fit to his or her individual time-course data was used to compute a starting and final WLI value on the basis of the total number of words heard during familiarization (range in number of words presented = 200–439). Conceptually, the final WLI value represents each individual's final end point of learning (i.e., reflecting the level of knowledge achieved by the end of the familiarization period) and, critically, is predicted to correlate with behavioral-learning outcomes. This computed final WLI value was used in the correlational analysis with the behavioral-test measure, as described in detail below.

Behavioral-data analyses

We compared infants' looking times to words from the artificial language with their looking times to nonword foils. For each infant, a cumulative looking-time score was computed per trial, excluding periods during which the infants' gaze was directed away from the screen before the trial offset. Because only a small subset of infants (~30%) completed Blocks 3 and 4, data were truncated to include only the first two blocks. Looking-time data were analyzed with a linear mixed-effects model including condition (word vs. nonword) and block as fixed effects and block-by-subject as a random slope. The analysis was implemented using R (Version 3.6.2; R Core Team, 2013) and the *lme4* package (Bates, Mächler, Bolker, & Walker, 2015). To establish the significance of the effect of the condition, we obtained p values from likelihood-ratio tests comparing the full model with the null model without the effect of interest (Winter, 2013).

Correlation between neural entrainment (familiarization phase) and preferential looking time (test phase)

To test the hypothesis that the WLI predicts infants' behavioral discrimination of words from foils, we calculated the correlation between (a) the difference between looking-time preferences to words and nonwords and (b) the individual log-fitted final WLI values (representing each infant's final sensitivity to the word structure after familiarization, based on the individual's logarithmic learning curve). We transformed the looking-time scores to reflect infants' preferential listening to nonwords over words while accounting for each individual's baseline looking time—proportional preference score = mean $LT_{nonwords}$/(mean $LT_{nonwords}$ + mean LT_{words}), where LT stands for looking time (Lany, Shoaib, Thompson, & Estes, 2018). A proportional score greater than .5 was indicative of longer looking to nonwords (novelty preference), and a score below .5 indicated longer looking to words (familiarity preference).

Comparison of neural entrainment in infants and adults (familiarization phase)

To compare statistical-learning trajectories in infants and adults, we analyzed data previously collected in adults ($N = 44$) under similar learning and testing conditions (Batterink & Paller, 2017). Briefly, EEG was recorded while adults were passively exposed to 12 min of an artificial language made up of four trisyllabic words, as in the present study. After familiarization, adults completed both an explicit rating task and a reaction time target-detection task (see the Supplemental Material for further details). We equated the language exposure between adults and infants by using an individual yoking procedure. For each adult, data from only the initial part of the 12-min familiarization period were included in the analysis, with this exact duration matched to one of the 25 infants. We applied the same data-analysis pipeline as the one described for the infant data. Possibly because of differences in skull size, adults showed a more focal distribution in both word-level and syllable-level entrainment than infants. Thus, for ITC and WLI computations in adults, we used a subset of nine frontocentral electrodes (F1, Fz, F2, FC1, C1, FCz, Cz, FC2, and C2) to avoid diluting our effects of interest (see Fig. S3 in the Supplemental Material for the distribution of the entrainment effect in adults).

To compare the time course of statistical learning in infants and adults, we used an independent-samples *t* test to determine whether *b* differed between the two groups, with significantly different *b* values taken to indicate that infants and adults show differences in statistical-learning rates.

Validation of analysis procedure in adults

Because data collected in infants and adults were subjected to identical data-analysis procedures, we tested the robustness of our infant-analysis pipeline by independently validating these analyses in data collected in adults. All validation checks were confirmed; see the Supplemental Material for further details.

Results

Familiarization phase

Overall neural entrainment. Similar to prior results in adults (Batterink & Paller, 2017), the current findings revealed that neural entrainment in infants across the familiarization period was characterized by clear peaks at the word and syllable frequencies but no clear peak at the bisyllable frequency (corresponding to 1.67 Hz). The ITC distribution at the word and syllable frequencies was relatively widespread across the scalp (Fig. 2a).

Neural-entrainment time course. Results were consistent with our predictions: WLI increased significantly as a function of exposure, providing evidence of on-line statistical learning of the underlying word units. Individual time-course data were modeled logarithmically because a logarithmic curve fitted the average group-level data better than a linear curve (log R^2 = .25, linear R^2 = .086). The increase in WLI over familiarization was reflected by a significant positive *b* value, representing a logarithmic increase, $y = a + b \times \log(x)$, mean b = 0.027, SE = 0.012, $t(24)$ = 2.16, p = .041, Cohen's d = 0.43. The mean modeled WLI intercept was 0.87 (SE = 0.074), and the mean modeled final WLI value was 1.04 (SE = 0.025). The observed *b* value of 0.027 corresponded to the 99th percentile in the distribution of *b* values to shuffled surrogate bundles, indicating that the WLI increase in the experimental data was unlikely to have occurred by chance and was not a general result of the analysis procedure. This result provides evidence that rapid word-level entrainment is a signature of statistical learning.

Figure 2b shows the logarithmic learning curve modeled on group-averaged data in infants (red). The increase in WLI was most pronounced over approximately the first 10 bundles, corresponding to roughly the first 100 word presentations, or about 90 s of exposure (Fig. 2b).

For comparison, the same analysis was conducted with $ITC_{syllable}$ as the dependent variable. Consistent with our previous finding that syllable-level entrainment in adults decreases over the course of exposure (Batterink & Paller, 2017, 2019), the current results showed that $ITC_{syllable}$ in the average group-level data decreased logarithmically (log R^2 = .48, linear R^2 = .21). Across infants' individual learning curves, the decrease in $ITC_{syllable}$ over familiarization was reflected by a significant negative *b* value, $y = a + b \times \log(x)$, mean b = −0.012, SE = 0.004, $t(24)$ = −2.90, p = .008, Cohen's d = 0.58.

Test phase

Results in the test phase were consistent with our predictions: Infants' looking times were significantly longer for nonword trials compared with word trials, reflecting a novelty preference (nonword: M = 13.87 s, SE = 1.22; word: M = 11.74 s, SE = 0.94; Fig. 3a). This difference was significant, as determined via a likelihood-ratio test—condition: $\chi^2(1)$ = 4.27, p = .039 (see Table S2 in the Supplemental Material for further details).

Correlation between neural entrainment (familiarization phase) and preferential looking time (test phase)

Infants' proportional preference score was significantly positively correlated with the modeled final WLI value, which we hypothesize reflects each infant's ultimate learning attainment (r = .54, n = 18, p = .022). We conducted a leave-one-out analysis after standard diagnoses for influential points, using an inspection of the individual beta values and Cook's distance (see the Supplemental Material for details). When the data were reanalyzed without a single identified influential point, the overall conclusion was maintained, but the effect became stronger (r = .73, n = 17, p < .001; Fig. 3b).

We further evaluated the relationship between the final WLI value and the proportional preference score by including two additional variables that could potentially influence infants' looking-time preference in the model (the length of familiarization and the number of completed test trials; n = 17). In this model, final WLI

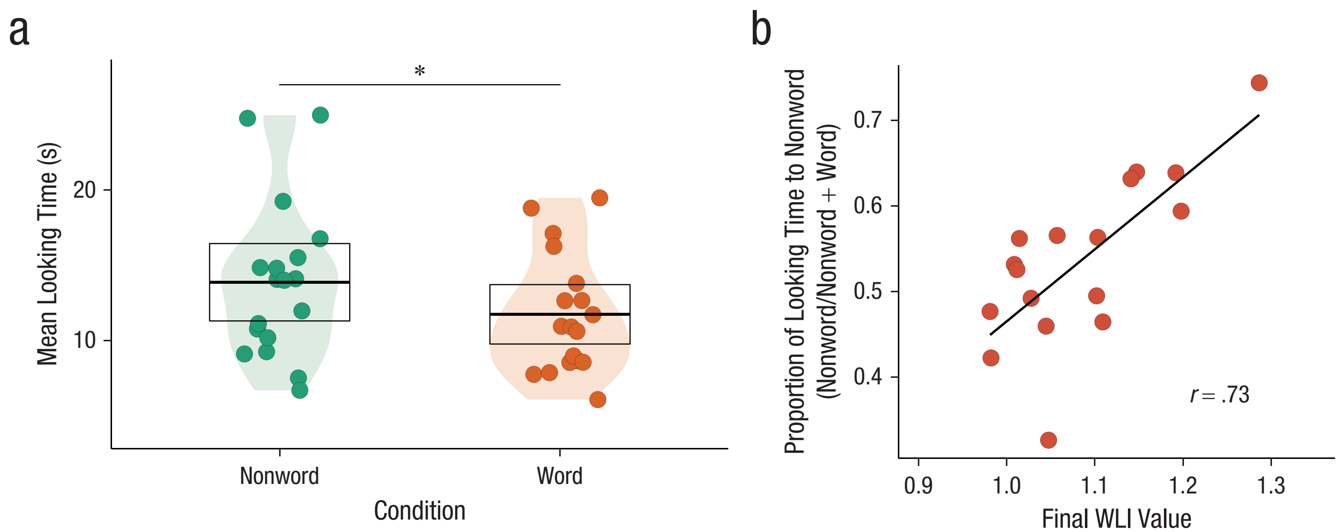

Fig. 3. Infants' looking-time results. Mean looking time (a) is shown for words and nonwords. Dots represent individual data points, horizontal bars represent group means, the height of the rectangles around the means indicates standard errors, and the width of the shaded areas indicates the density of the data. The asterisk indicates a significant difference between the means for words and nonwords. The scatterplot (b; with best-fitting regression line) shows the correlation between modeled final word-learning index (WLI) value (representing relative neural entrainment to words during the familiarization phase) and proportional looking-time measure at test (r = .73, n = 17, p < .001; influential point excluded).

significantly predicted the proportional preference score at test (b = 0.71, SE = 0.014), t(15) = 3.0, p = .011, whereas length of familiarization (b = −0.00019, SE = 0.00013), t(15) = −1.48, p = .16, and number of completed test trials (b = 0.018, SE = 0.014), t(15) = 1.34, p = .20, did not.

Comparison of neural entrainment in infants and adults

Similar to infants' results, in adults, a logarithmic curve better fitted the data than did a linear curve (log R^2 = .30, linear R^2 = .17). Individual learning curves in infants and adults did not show significantly different WLI slopes—mean slope in adults = 0.028, SE = 0.010, t(67) = −0.087, p = .93, Cohen's d = 0.022 (see Table S5 in the Supplemental Material; Fig. 2b). To quantify evidence for the alternative hypothesis that the slopes differ between infants and adults versus the null hypothesis that the slopes do not differ between the two groups, we computed a Bayes factor (BF). The BF revealed that the observed data were more than 5 times more likely to arise under the null hypothesis compared with the alternative hypothesis (BF_{01} = 5.27). This is considered moderate (Wagenmakers et al., 2018) or substantial (Jarosz & Wiley, 2014) evidence for the null hypothesis of no difference between groups. These results suggest that infants and adults became sensitive to the statistical structure of the speech stream at similar rates. We note that WLI values were numerically higher in adults than in infants across the familiarization period (Fig. 2b; see also Fig. S3 in the Supplemental Material), though we restricted our statistical comparisons to the learning trajectories between the groups. A direct comparison of the overall entrainment values between groups is not straightforward, as anatomical differences and age-related changes in EEG spectral power (both across and between frequency bands) may influence neural entrainment effects between groups, independently of learning-specific effects. In summary, there was no evidence of statistically different learning trajectories between infants and adults (see Table S5 in the Supplemental Material).

Figure 2b shows the logarithmic learning curves modeled on group-averaged data for infants and adults. This analysis produced roughly similar estimated parameters between groups (infants: b = 0.014, modeled initial WLI value = 0.95, modeled final WLI value = 1.05; adults: b = 0.025, modeled initial WLI value = 0.94, modeled final WLI value = 1.12). Again, WLI values were numerically higher in adults compared with infants.

Finally, as a comparison, the same analysis was conducted on the adult data with $ITC_{syllable}$ as the dependent variable. $ITC_{syllable}$ in the average group-level data decreased numerically over familiarization, though it did not follow a clear logarithmic function (log R^2 = .17, linear R^2 = .15). $ITC_{syllable}$ across individual learning curves showed a marginally significant decrease when modeled both linearly (p = .065) and logarithmically (p = .087).

Discussion

Summary of results

These results provide evidence that preverbal infants rapidly attune to the statistical structure of an artificial language. During the 6-min familiarization period, we observed an increase in the WLI, the ratio of neural entrainment to the word rate relative to the syllable rate. We interpret this logarithmic increase as reflecting a shift in perception from isolated syllables to coherent wordlike units. In the test phase, infants showed a preference for nonword foils compared with words, and crucially, infants' WLI scores during familiarization significantly correlated with their preference for nonwords during the test phase. Finally, the overall pattern of neural entrainment and the rate of WLI progression in infants was comparable with the effects observed in adult learners (Batterink & Paller, 2017).

New insights into statistical learning in infants

To our knowledge, this research is the first to shed light on the time course and dynamics of statistical learning in preverbal infants. The WLI time-course data were best modeled using a logarithmic function rather than a linear one. WLI values showed the most rapid increase within approximately the first 100 word presentations, or about 90 s of familiarization (Fig. 2b). These findings support a number of previous studies of statistical learning in infants showing successful learning after only 2 min of familiarization (e.g., Saffran et al., 1996; Saffran, 2001). At least within the context of this artificial-speech-segmentation paradigm, our results suggest that 90 s to 2 min may represent a sweet spot for learning, with infants showing only incremental gains after this point. Our finding of a logarithmic learning curve converges with previous findings from statistical-learning work in adults (Siegelman et al., 2018) and across other learning domains (Dayan & Cohen, 2011; Karni & Sagi, 1993). One possibility is that infants' attention may begin to wane after 2 min, so additional exposure is of limited effectiveness. Another possibility is that the neural processes responsible for discovering the statistics converge into a pattern within the first few minutes of exposure. Additional exposure may then solidify and strengthen this initial discovery, rather than producing novel discoveries of the underlying structure that would lead to additional gains in learning.

We also report the first evidence in infants linking neural mechanisms engaged during statistical learning to subsequent behavioral outcomes. We observed a significant positive correlation between infants' WLI score during familiarization and their preference for nonwords during the test phase. Along with previous results in adults (Batterink & Paller, 2017, 2019), this finding provides additional support for an association between proportional phase locking of ongoing neural oscillations to the word frequency and the successful discovery and encoding of words in continuous speech. Infants with stronger entrainment to the word frequency relative to the syllable frequency were more likely to prefer nonwords—that is, novel not-yet-learned structures. Conversely, infants with lower WLI scores allocated more attention to the familiar words. We propose that infants with higher WLI scores were more efficient learners and thus were more likely to show increased attention toward the novel structures relative to the words acquired by the end of familiarization. Infants with lower WLI may have been less efficient learners and thus showed a familiarity preference as they continued to acquire information about those structures. This interpretation is consistent with the processing-load account, which suggests that infants are more likely to show a familiarity preference when stimuli are relatively complex and difficult to process and to show a novelty preference when stimuli are simpler and easier to process (Hunter, Ames, & Koopman, 1983). Although the positive correlation between the WLI and the behavioral novelty preference is in line with theoretical predictions and previous findings in adults, it will be important to replicate this finding in future studies, given the relatively low sample size in this analysis (n = 18).

Our neural-entrainment time-course results suggest that infants and adults follow similar learning trajectories when tracking probability information among speech sounds. We found no evidence of significant differences between groups on either time-course values or overall WLI values, and a Bayesian approach indicated substantial evidence for the null hypothesis of no difference in the learning trajectory between groups. These results suggest an equally robust and efficient statistical-learning mechanism throughout the life span, supporting previous work (e.g., Saffran, Newport, Aslin, Tunick, & Barrueco, 1997; Thiessen, Girard, & Erickson, 2016). If statistical-learning ability were considered in isolation, critical-period effects and differences in the acquisition of first and second languages would be unexplained. However, other aspects of language learning that are known to be more dependent on age, such as acquisition of phonetic categories and syntactic dependencies, may impact statistical learning. Previous studies suggest that statistical learning interacts with prior linguistic knowledge and that variance in statistical-learning performance can be predicted by the resemblance of stimuli to participants' native language (Siegelman et al., 2018).

The current findings raise the theoretical possibility that neural entrainment may represent a common mechanism underlying statistical learning in both infants and adults. In both age groups, relative neural entrainment to words correlated with postfamiliarization behavioral measures of statistical learning. Although these results cannot directly speak to causality between neural entrainment and statistical learning, there is theoretical grounding for the idea that neural entrainment to structure may contribute to statistical learning. Synchrony between neural oscillations and external stimuli allows for phases of high neural excitability to align with stimulus events, enabling more efficient processing (Thut, Miniussi, & Gross, 2012). In the context of statistical learning, aligning phases of neural excitability to an ongoing temporal pattern may facilitate processing of predictable stimuli (Schroeder & Lakatos, 2009), promoting the emergence of cohesive word percepts. Consistent with the idea that neural entrainment plays a causal role in processing, recent findings from studies of speech perception have demonstrated that enhancing neural entrainment to the auditory envelope of speech, via electrical brain stimulation, improves speech intelligibility (Riecke, Formisano, Sorger, Baskent, & Gaudrain, 2018; Wilsch, Neuling, Obleser, & Herrman, 2018). These results provide additional evidence that stronger neural entrainment to relevant features of the acoustic signal leads to more effective auditory processing.

An alternative account to the notion that neural entrainment to the auditory stream facilitates statistical learning is that our measure of entrainment may simply reflect downstream effects of the learning process. For example, neural entrainment in this context may reflect participants' perception of the speech stream (for evidence of neural entrainment reflecting perception, see Ding et al., 2016; Nozaradan et al., 2011), which may merely follow from the mechanism responsible for identifying the most relevant or parsimonious units in an input stream. The current design also does not allow us to disentangle whether the WLI reflects automatic rhythmic entrainment to trisyllables, rather than word-level entrainment, given that all words in the stream were the same length. However, in a recent study in adults, we also found phase locking to words of variable durations embedded in a continuous speech stream, which predicted subsequent word knowledge (Batterink, 2020). This result suggests that the neural data in the present study at least partially reflect specific word-level entrainment, rather than general rhythmic entrainment to the trisyllabic stream. Lastly, we tested infants' discrimination of words versus nonwords, rather than words versus partial words (aka *partwords*). In principle, it is possible that infants show different preferences between words (previously encountered trisyllabic structures) and nonwords (trisyllables never encountered in that linear order) by relying on familiarity alone, without having segmented the underlying statistically defined constituents in the speech stream. However, prior behavioral studies that controlled for the effect of bigram frequency showed that infants also succeeded in discriminating words from partwords (Saffran et al., 1996) and from frequency-balanced partwords (Aslin, Saffran, & Newport, 1998), providing evidence that infants successfully learn the underlying units in the speech stream (see the Supplemental Material for further discussion of this issue). Our results further support the hypothesis that infants' discrimination of words from nonwords reflects successful segmentation of the speech stream, as opposed to simple recognition of previously encountered strings, by demonstrating that on-line neural chunking of the stream into its statistically defined constituents is tightly related to subsequent word and nonword preferences.

To conclude, statistical learning has been proposed to be an important mechanism for both first- and second-language acquisition. The demonstration of similar on-line statistical learning in both adults and infants provides support for this idea, suggesting an early emerging cognitive mechanism that is maintained throughout the life span.

Transparency

Action Editor: Rebecca Treiman

Editor: D. Stephen Lindsay

Author Contributions

D. Choi and L. J. Batterink contributed equally to this article. D. Choi, L. J. Batterink, A. K. Black, and J. F. Werker developed and designed the study. D. Choi and L. J. Batterink programmed the experiment and implemented the study. Infant testing and data collection were performed by D. Choi and A. K. Black. L. J. Batterink analyzed and interpreted the electroencephalogram (EEG) data. D. Choi and A. K. Black analyzed and interpreted the behavioral data. D. Choi analyzed the EEG–behavior correlation. D. Choi, L. J. Batterink, and A. K. Black drafted the manuscript. K. A. Paller and J. F. Werker provided revisions and critical feedback on the manuscript. All of the authors approved the final manuscript for publication.

Declaration of Conflicting Interests

The author(s) declared that there were no conflicts of interest with respect to the authorship or the publication of this article.

Funding

This research was funded by grants to J. F. Werker from the Natural Sciences and Engineering Research Council of Canada (RGPIN-2015-03967) and the Canada Foundation for Innovation John R. Evans Leaders Fund (33096).

Open Practices

All data and stimuli for this study have been made publicly available via OSF and can be accessed at https://osf.io/6a9je/. The sample size, design, and analysis plans were preregistered at https://osf.io/wamv4. Only the preregistered electroencephalogram (EEG) analyses are reported here; the preregistration states that event-related potential (ERP) analyses would also be conducted, but we did not obtain sufficient data and thus could not run the ERP analyses. Several other minor deviations from the preregistration (related to the power analysis, the EEG entrainment analyses, and the method used to analyze looking-time scores) are detailed in the Supplemental Material available online. The complete Open Practices Disclosure for this article can be found at http://journals.sagepub.com/doi/suppl/10.1177/0956797620933237. This article has received the badges for Open Data, Open Materials, and Preregistration. More information about the Open Practices badges can be found at http://www.psychologicalscience.org/publications/badges.

ORCID iDs

Dawoon Choi https://orcid.org/0000-0002-3946-9901
Ken A. Paller https://orcid.org/0000-0003-4415-4143

Acknowledgments

We thank Sav Nijeboer for her support in recruiting infants.

Supplemental Material

Additional supporting information can be found at http://journals.sagepub.com/doi/suppl/10.1177/0956797620933237

References

Aslin, R. N., Saffran, J. R., & Newport, E. L. (1998). Computation of conditional probability statistics by 8-month-old infants. *Psychological Science*, *9*, 321–324.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1). doi:10.18637/jss.v067.i01

Batterink, L. (2020). Syllables in sync form a link: Neural phase-locking reflects word knowledge during language learning. *Journal of Cognitive Neuroscience*, *32*, 1735–1748.

Batterink, L. J., & Paller, K. A. (2017). Online neural monitoring of statistical learning. *Cortex*, *90*, 31–45.

Batterink, L. J., & Paller, K. A. (2019). Statistical learning of speech regularities can occur outside the focus of attention. *Cortex*, *115*, 56–71. doi:10.1016/j.cortex.2019.01.013

Bergelson, E., & Swingley, D. (2012). At 6–9 months, human infants know the meanings of many common nouns. *Proceedings of the National Academy of Sciences, USA*, *109*, 3253–3258. doi:10.1073/pnas.1113380109

Buiatti, M., Peña, M., & Dehaene-Lambertz, G. (2009). Investigating the neural correlates of continuous speech computation with frequency-tagged neuroelectric responses. *NeuroImage*, *44*, 509–519.

Cirelli, L. K., Spinelli, C., Nozaradan, S., & Trainor, L. J. (2016). Measuring neural entrainment to beat and meter in infants: Effects of music background. *Frontiers in Neuroscience*, *10*, Article 229. doi:10.3389/fnins.2016.00229

Dayan, E., & Cohen, L. G. (2011). Neuroplasticity subserving motor skill learning. *Neuron*, *72*, 443–454.

Dehaene-Lambertz, G., & Gliga, T. (2004). Common neural basis for phoneme processing in infants and adults. *Journal of Cognitive Neuroscience*, *16*, 1375–1387.

Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*, 9–21.

Ding, N., Melloni, L., Zhang, H., Tian, X., & Poeppel, D. (2016). Cortical tracking of hierarchical linguistic structures in connected speech. *Nature Neuroscience*, *19*, 158–164.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*, 175–191.

Fló, A., Brusini, P., Macagno, F., Nespor, M., Mehler, J., & Ferry, A. L. (2019). Newborns are sensitive to multiple cues for word segmentation in continuous speech. *Developmental Science*, *22*(4), Article e12802. doi:10.1111/desc.12802

Fujioka, T., Mourad, N., He, C., & Trainor, L. J. (2011). Comparison of artifact correction methods for infant EEG applied to extraction of event-related potential signals. *Clinical Neurophysiology*, *122*, 43–51.

Graf Estes, K., Evans, J. L., Alibali, M. W., & Saffran, J. R. (2007). Can infants map meaning to newly segmented words? Statistical segmentation and word learning. *Psychological Science*, *18*, 254–260.

Graf Estes, K., & Lew-Williams, C. (2015). Listening through voices: Infant statistical word segmentation across multiple speakers. *Developmental Psychology*, *51*, 1517–1528. doi:10.1037/a0039725

Hunter, M. A., Ames, E. W., & Koopman, R. (1983). Effects of stimulus complexity and familiarization time on infant preferences for novel and familiar stimuli. *Developmental Psychology*, *19*, 338–352.

Jarosz, A. F., & Wiley, J. (2014). What are the odds? A practical guide to computing and reporting Bayes factors. *The Journal of Problem Solving*, *7*(1), Article 2. doi:10.7771/1932-6246.1167

Kabdebon, C., Pena, M., Buiatti, M., & Dehaene-Lambertz, G. (2015). Electrophysiological evidence of statistical learning of long-distance dependencies in 8-month-old preterm and full-term infants. *Brain and Language*, *148*, 25–36.

Karni, A., & Sagi, D. (1993). The time course of learning a visual skill. *Nature*, *365*, 250–252.

Klatt, D. H. (1980). Speech perception: A model of acoustic-phonetic analysis and lexical access. In R. A. Cole (Ed.), *Perception and production of fluent speech* (pp. 243–288). Mahwah, NJ: Erlbaum.

Lany, J., Shoaib, A., Thompson, A., & Estes, K. G. (2018). Infant statistical-learning ability is related to real-time language processing. *Journal of Child Language*, *45*, 368–391.

Mourad, N., Reilly, J. P., de Bruin, H., Hasey, G., & MacCrimmon, D. (2007). A simple and fast algorithm for automatic suppression of high-amplitude artifacts in EEG data. In *2007 IEEE International Conference on Acoustics, Speech and Signal Processing - ICASSP '07* (pp. 1393–1396). Retrieved from https://doi.org/10.1109/ICASSP.2007.366699

Nozaradan, S., Peretz, I., Missal, M., & Mouraux, A. (2011). Tagging the neuronal entrainment to beat and meter. *The Journal of Neuroscience*, *31*, 10234–10240.

R Core Team. (2013). R: A language and environment for statistical computing [Computer software]. Retrieved from http://www.R-project.org

Riecke, L., Formisano, E., Sorger, B., Baskent, D., & Gaudrain, E. (2018). Neural entrainment to speech modulates speech intelligibility. *Current Biology*, *28*, 161–169.

Saffran, J. R. (2001). Words in a sea of sounds: The output of infant statistical learning. *Cognition*, *81*, 149–169.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, *274*, 1926–1928.

Saffran, J. R., & Kirkham, N. Z. (2018). Infant statistical learning. *Annual Review of Psychology*, *69*, 181–203.

Saffran, J. R., Newport, E. L., Aslin, R. N., Tunick, R. A., & Barrueco, S. (1997). Incidental language learning: Listening (and learning) out of the corner of your ear. *Psychological Science*, *8*, 101–105.

Schroeder, C. E., & Lakatos, P. (2009). Low-frequency neuronal oscillations as instruments of sensory selection. *Trends in Neurosciences*, *32*(1), 9–18.

Shi, R., & Werker, J. F. (2001). Six-month-old infants' preference for lexical words. *Psychological Science*, *12*, 70–75.

Siegelman, N., Bogaerts, L., Kronenfeld, O., & Frost, R. (2018). Redefining "learning" in statistical learning: What does an online measure reveal about the assimilation of visual regularities? *Cognitive Science*, *42*, 692–727.

Swingley, D. (2005). 11-month-olds' knowledge of how familiar words sound. *Developmental Science*, *8*, 432–443.

Teinonen, T., Fellman, V., Näätänen, R., Alku, P., & Huotilainen, M. (2009). Statistical language learning in neonates revealed by event-related brain potentials. *BMC Neuroscience*, *10*, Article 21. doi:10.1186/1471-2202-10-21

Thiessen, E. D., Girard, S., & Erickson, L. C. (2016). Statistical learning and the critical period: How a continuous learning mechanism can give rise to discontinuous learning. *Wiley Interdisciplinary Reviews: Cognitive Science*, *7*, 276–288.

Thut, G., Miniussi, C., & Gross, J. (2012). The functional importance of rhythmic activity in the brain. *Current Biology*, *22*, R658–R663.

Wagenmakers, E.-J., Love, J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., . . . Morey, R. D. (2018). Bayesian inference for psychology. Part II: Example applications with JASP. *Psychonomic Bulletin and Review*, *25*, 58–76.

Wilsch, A., Neuling, T., Obleser, J., & Herrman, C. S. (2018). Transcranial alternating current stimulation with speech envelopes modulates speech comprehension. *NeuroImage*, *172*, 766–774.

Winter, B. (2013). Linear models and linear mixed effects models in R with linguistic applications. *arXiv*. Retrieved from https://arxiv.org/abs/1308.5499