NeuroImage 281 (2023) 120384

Contents lists available at ScienceDirect

NeuroImage

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

Auditory learning of recurrent tone sequences is present in the newborn's brain

Brigitta Tóth [a,*,1], Péter Kristóf Velősy [a,b,1], Petra Kovács [a,b], Gábor Peter Háden [a,c], Silvia Polver [d], Istvan Sziller [e], István Winkler [a]

[a] Institute of Cognitive Neuroscience and Psychology, Research Centre for Natural Sciences, Budapest, Hungary
[b] Department of Cognitive Science, Budapest University of Technology and Economics, Budapest, Hungary
[c] Department of Telecommunications and Media Informatics, Faculty of Electrical Engineering and Informatics, Budapest University of Technology and Economics, Budapest, Hungary
[d] Department of Developmental Psychology and Socialisation, University of Padova, Padova, Italy
[e] Division of Obstetrics and Gynecology, DBC - Szent Imre University Teaching Hospital, Budapest, Hungary

ARTICLE INFO

Keywords:
Auditory processing
Pattern repetition
Newborn infants
Event-related potentials (ERP)

ABSTRACT

The seemingly effortless ability of our auditory system to rapidly detect new events in a dynamic environment is crucial for survival. Whether the underlying brain processes are innate is unknown. To answer this question, electroencephalography was recorded while regularly patterned (REG) versus random (RAND) tone sequences were presented to sleeping neonates. Regular relative to random sequences elicited differential neural responses after only a single repetition of the pattern indicating the existence of an innate capacity of the auditory system to detect auditory sequential regularities. We show that the newborn auditory system accumulates evidence only somewhat longer than the minimum amount determined by the ideal Bayesian observer model (the prediction from a variable-order Markov chain model) before detecting a repeating pattern. Thus, newborns can quickly form representations for regular features of the sound input, preparing the way for learning the contingencies of the environment.

1. Introduction

A large body of evidence suggests that the adult human brain tracks temporal/sequential regularities in the auditory input, establishing and continuously refining internal models that generate predictions for future events (Friston and Kiebel, 2009; Rubin et al., 2016; Winkler et al., 2009; McDermott et al., 2013; Paavilainen, 2013; Saffran et al., 1999). So far, most evidence has been obtained from a paradigm in which an incoming sound was inconsistent with these predictions, and a 'surprise' response termed mismatch negativity (MMN) could be observed (Näätänen et al., 1978; for a recent review, see Fitzgerald and Todd, 2020; Friston, 2005). The neonate's brain also responds to sounds violating some regularity of the preceding sound sequence (termed the mismatch response [MMR]; Alho et al., 1990; Kushnerenko et al., 2013). However, in infants, the temporal dynamics and the brain areas involved in discovering complex temporal patterns are yet unknown. Here we measure for the first time the processes of learning recurrent patterns within a tone sequence by measuring when the newborn's brain detects the regularity of a tone sequence and comparing it to the expected performance of an ideal Bayesian observer.

Behaviorally relevant sound sequences often have regular auditory features (e.g., a series of footsteps, etc.). Detection of regular and repetitive auditory features allows the listeners to recognize objects and to predict how they will behave soon (Bendixen, 2014; Schröger et al., 2014; Winkler and Schröger, 2015). Much of the evidence supporting this notion is indirect since based on studies measuring the MMN response with electro- and magnetoencephalography (Näätänen et al., 1978); for a recent review, see Fitzgerald and Todd, 2020. In the MMN-eliciting paradigms, regular sounds (standard) are occasionally interrupted by 'deviant' sounds, which do not confirm the regularity of the preceding standards. MMN is typically assessed by comparing the response elicited by deviants with that measured for the standards or control sounds matching the acoustic parameters of the deviant but being regular within their context (Heilbron and Chait, 2018; Jacobsen

---

* Corresponding author.
E-mail address: toth.brigitta@ttk.hu (B. Tóth).

[1] These authors contributed equally.

https://doi.org/10.1016/j.neuroimage.2023.120384
Received 19 May 2023; Received in revised form 13 August 2023; Accepted 19 September 2023

et al., 2003; for a review; see Sussman et al., 2008). Studies measuring the auditory MMN suggest that the human brain is sensitive to regularities over multiple timescales and complexity (Paavilainen et al., 2007; Barascud et al., 2014; Bendixen et al., 2012; Bekinschtein et al., 2009). While sound repetition is considered a simple regularity, MMN is also elicited by violating more complex regularities, such as feature trends, repeating sound patterns, inter-sound feature relationships (including that between non-adjacent sounds), variation in the regular sound features, etc. Violating any of these complex regularities elicits the MMR also in infants (Basirat et al., 2014; Háden et al., 2015; Virtala et al., 2013), suggesting that infants also detect and represent many of the regular aspects of the acoustic input.

A significant limitation is that studies using regularity violation (MMN response) only provide indirect evidence that the brain has detected some statistical regularity. Filling this gap, Barascud and Collages (2016) targeted the brain processes subserving the learning of recurrent patterns within a tone sequence. Tone sequences presented in a regular complex recurrent pattern elicited neural responses in primary sensory areas, the hippocampus, and the inferior frontal gyrus (IFG), even when these were not behaviorally relevant. The temporal

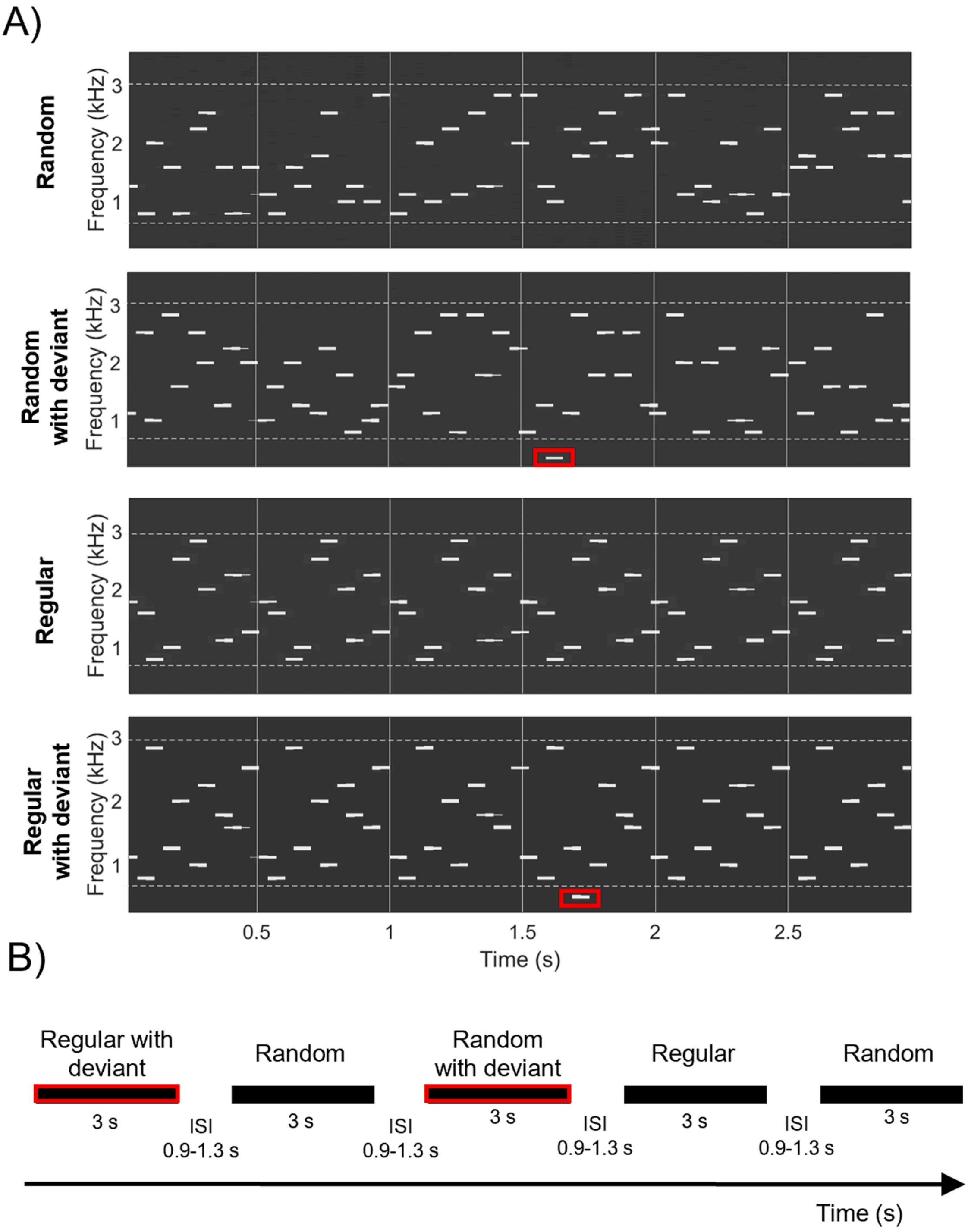

Fig. 1. A) Stimuli with and without a repeating pattern (Regular: REG; Random: RAND) and with or without a frequency deviant (0.25 probability within the experiment. each). The *y*-axis denotes frequency; the *x*-axis shows time. White rectangles represent tones, red frame highlights deviant tones. Green dashed lines show the borders between cycles of the pattern repeated in the REG sequences. B) An example sequence of the four types of stimulus.

properties of listeners' behavioral and brain responses were then compared to a Bayesian ideal observer model (the probabilistic predictions of a variable-order Markov model). Since the model was unconstrained by memory and attention, it set the limit for the most efficient perceptual recognition possible. The results demonstrated that dynamic changes in the neural responses correlated with informational entropy (predictability of ongoing auditory input). Another recent study from Southwell and Chait (2018) showed that EEG responses are sensitive to the predictability of randomly generated tone-pip sequences. For instance, higher-amplitude MMN responses were elicited by the same deviant tone when presented in the context of a structural regularity than within a random tone cloud. In addition to the commonly observed deviant detecting fronto-temporal network, the regularity effect was underpinned by sources in the right temporal pole and the orbitofrontal cortex. Thus, these findings indicate that the adult brain keeps track of the patterns within the auditory input. However, whether the processes of tracking and representing patterns within a sound sequence are innate is still unknown.

The present study aims to test how the newborn brain detects patterns within a sound sequence. Newborn infants were presented with a series of tones with and without regular patterns (termed regular [REG] or random [RAND]; see Southwell and Chait, 2018) while recording high-density EEG. Successive tones (differing only in frequency) were highly predictable in the REG while unpredictable in the RAND sequences. Since unique tone sequences were created for each trial, for detecting the pattern repetition, the newborn brain had to build up a new internal model of the regular pattern on each trial to detect when it was violated (Fig. 1). Based on predictive frameworks (e.g., Friston and Kiebel, 2009) suggesting that the brain compares incoming sensory information against internal models established by the statistics of the preceding sensory input, we expect regular tones (REG) to elicit different neural activity than RAND sequences commencing when pattern repetition is detected. We also expect differences between the MMR responses elicited by deviant tones appearing within the two types of sequences because violating the standard range of frequencies is common in both conditions (Winkler et al., 1990); pattern violation can only happen in the REG condition. The accumulation of stimulus statistics in the newborn brain (as measured by EEG) is then compared to predictions of a variable-order Markov chain model (developed by Barascud et al., 2016), which assesses the plateau performance of the human perceptual system.

2. Materials and methods

2.1. Participants

Thirty-three healthy full-term newborn infants (0–4 days of age, 17 male, APGAR score 9/10 or above) were measured. They had a mean gestational age of 43.24 weeks (SD = 2.39) and a mean birthweight of 3857.28 g (SD = 232.62). All newborns were firstborns, none twins, had normal hearing and passed the Brainstem Evoked Response Audiometry (BERA) test. Informed consent was obtained from one or both parents, and the infant's mother could opt to be present during the recording. The study fully complied with the World Medical Association Helsinki Declaration and all applicable national laws and was approved by the National Public Health Center of Hungary (TUKEB).

2.2. Stimuli

The stimuli used in this experiment are illustrated in Fig. 1. They were identical to the study of Southwell and Chait (2018). 3000 ms long tone sequences were created from 50 ms long pure tones taken from the 198 to 3563 Hz frequency range (26 frequency values in logarithmic spacing, 12 % increase per step). Each sequence contained tones of ten unique frequency components. The tones were organized in a regular (REG) or random (RAND) sequential arrangement. In REG sequences, the tones were arranged in a 10-tone pattern and presented six times (60 tones of 50 ms each amounting to the 3000 ms stimuli; one cycle is 500 ms long). Corresponding RAND sequences were generated for each REG sequence by randomizing the order of the same tones with the caveat that two tones of the same frequency could not follow each other. To ensure that each sequence contained roughly the same bandwidth of frequency components, a random subset of 13 adjacent tones was chosen, then 10 of these tones were selected randomly to create the stimulus.

Half of the sequences contained a frequency deviant to determine whether the deviant response reflects only the detection of an auditory feature change (i.e., a frequency outside the range of the rest of the frequencies) or is also influenced by the sequence context (the regularity of the sequence). To this end, half of the RAND and REG trials had one of the 60 tones replaced by a frequency deviant, either higher or lower (50–50 %) than the rest of the tones in the sequence by at least 24 %. The deviant followed the onset of the stimulus by 1500–2750 ms, which means that a minimum of three cycles had been completed within REG stimuli before the presentation of a deviant. Based on the work of Barascud et al. (2016) and Southwell et al. (2017), this placement ensures that the regularities in the REG stimuli have been sufficiently processed by an adult listener by the time the deviant appears. Deviation from a range of frequencies from which most of the tones of the sequence are taken elicits the MMN response in adults (Winkler et al., 1990).

Thus, the stimulus generation method yielded four types of sequences, each comprising the same ten frequency components. Two sequences contained no deviants – denoted REG-only and RAND-only, while two contained a deviant – deviant-REG and deviant-RAND. Each infant received uniquely generated stimuli on each trial. The four types of stimuli were delivered with equal probabilities (0.25) in a randomized order with the constraint that subsequent stimuli were of different conditions.

2.3. Procedure

EEG was recorded in a dedicated experimental room at the Department of Obstetrics-Gynecology, Szent Imre Hospital, Budapest. The newborns were asleep during the stimulus presentation. Infants' sleep state was determined based on behavioral criteria by Anders et al. (1971). Only infants in quiet sleep for the whole 5-minute duration were included in the study. In addition to the behavioral criteria employed, the EEG signal was visually inspected to ensure muscle tension was a tonic, respiration was regular, and large eye movements were absent.

Together 960 stimuli – 240 of each of the four conditions — were presented using a Maya 22 USB external soundcard and ER•2 Insert Earphones (Etymotic Research Inc., Elk Grove Village, IL, USA) placed into the infants' ears via ER2 Foam Infant Ear-tips. Tones were presented at a comfortable intensity of cca. 70 dB SPL. The presentation of the stimuli was implemented in the E-Prime stimulus presentation software (Psychology Software Tools, Inc., Pittsburgh, PA). EEG was recorded throughout the stimulus presentation. The inter-stimulus interval (ISI, offset to onset) was 900–1300 ms (random with even distribution, 1 ms step). The experiment took 45 min overall.

2.4. Analysis

2.4.1. EEG recording and preprocessing

BrainAmp DC amplifier with a 64-channel sponge-based electrode system (saltwater sponges and passive Ag/AgCl electrodes) and BrainVision Recorder were employed for EEG recording. The sampling rate was 500 Hz with a 100 Hz online low-pass filter applied. Electrodes were placed according to the International 10/20 system. The Cz channel served as the reference electrode while the ground electrode was placed on the midline of the forehead. During the recording, impedances were kept below 15 kΩ.

EEG signals were preprocessed in Matlab 2020a using the EEGlab toolbox version 2022.0 (Delorme et al., 2007). The recordings were low-pass filtered at 0.01 Hz and high-pass filtered at 30 Hz. All filtering was performed with a Hamming windowed sinc FIR filter of the order of 33,000 (pop_eegfiltnew in EEGlab). Recordings were later re-referenced to the average of all electrodes. A maximum of six malfunctioning EEG channels was allowed per participant. These were interpolated using the default spline interpolation algorithm implemented in EEGlab. Recordings were epoched separately for the REG-RAND comparisons and the deviant response analysis. On average, 1,6 (stdDev:0.9) bad electrodes were interpolated per participant per condition. No infants were excluded based on the criteria regarding the number of wrong channels.

2.4.2. REG and RAND event-related neural response analysis (time-locked to the sequence onset)

From the continuous EEG record, epochs were extracted between −500 and +3500 ms relative to the onset of the stimulus sequences, separately for REG and RAND stimuli; stimuli with (deviant-REG/RAND) and without a deviant (REG/RAND-only) were separately collapsed. Baseline correction was calculated from the 100 ms pre-stimulus interval. Artifacts were removed based on visual inspection: epoch was rejected from further analysis if physiological (eye movements, muscle artifacts) or external (i.e., line noise from electrical activity) artifacts were found. After artifact rejection, the epoch numbers were equalized across conditions, separately for each infant, by randomly removing epochs until the epoch number in both conditions matched that of the condition with the lower number of epochs for the given infant. As a result, an average of 162.28 epochs were retained for each infant and condition (σ: 23.28, min: 118, max: 203). The global field power (GFP) was estimated as the mean root square of the signal from all electrodes for each time point to compare the overall difference between REG and RAND trials. The GFP signal reflects the power of the evoked response at each time point (for comparability with Southwell and Chait, 2018).

For comparing the GFP of REG and RAND, a cluster-based statistic was employed to avoid a priori choices of time ranges of interest (Maris and Oostenveld, 2007). At each time point between 0 and 3000 msec from stimulus onset, individual averaged epochs were compared between the two conditions (RAND vs. REG), creating pairwise group contrasts matrices of t-values ($p < 0.05$). Two-tailed t-tests were computed on time clusters and controlled for family-wise error with a rate of 0.05 using 10,000 permutations to create the null model (Maris and Oostenveld, 2007).

2.4.3. Analysis of the deviant evoked response (time-locked to the onset of the deviant tone)

The time window to extract the epochs was −100 to 800 ms relative to the onset of the deviant tone. Each deviant epoch was randomly paired with a unique non-deviant trial of the same, the epoch from which was measured at the same time point relative to stimulus onset and the same onset and offset times. Artifact rejection and the equalization of the epoch numbers for the two conditions within each infant were identical to the RAND vs. REG stimulus-response analysis (see Section 2.4.2.). An average number of 161.34 deviant epochs were retained for each infant and condition (σ: 23.96, min: 118, max: 203). Event-related potentials were calculated for each condition separately for a fronto-central region of interest (F1, F3, Fz, F2, F4, FC3, FC1, FC2, FC4, C3, C1, C2). The same cluster-based permutation test was used as the above analysis to identify temporal clusters where the response to the deviant significantly differed from that to the non-deviant.

2.4.4. EEG source localization

EEG source reconstruction was performed with the Brainstorm toolbox (version 2022 January; Tadel et al., 2011). An age-appropriate template, based on MRI scans of two weeks old infants, was used in this study (O'Reilly et al., 2021) and, along with default electrode locations, were entered into the forward boundary element head model (BEM) provided by the openMEEG algorithm (Gramfort et al., 2011). For modeling the time-varying source signals (current density) for all cortical voxels, a minimum norm estimate inverse solution was applied using dynamical Statistical Parametric Mapping normalization (Dale et al., 2000), separately for RAND and REG trials (only responses to REG- and RAND-only stimuli were included). Averaging current density across voxels yielded time series for 62 cortical areas (region of interest ROIs), defined by the standardized parcellation scheme introduced by Desikan and Killiany (Desikan et al., 2006). The noise covariance was calculated from the 100 ms pre-stimulus interval. Source orientations were considered constrained (perpendicular to the surface), and source signals were reconstructed at 15,000 vertices describing the pial surface. (1.048–1.594 s; see Section 3.1). Current source density in RAND and REG conditions were calculated, and permutation-based (Monte Carlo method $N = 1000$) paired sample t-tests (alpha level = 0.01) was performed.

2.4.5. Modeling the detection of repeating patterns by the ideal Bayesian observer

To model the benchmark of the auditory expectation of the listeners, a variable-order Markov model was used to quantify the predictability of each tone within the stimuli (based on Barascud et al., 2016; Information Dynamics of Music [IDyOM], developed by Pearce 2005; Pearce and Wiggins, 2006, Pearce et al., 2010). IDyOM implements unsupervised statistical learning to acquire transitional dependencies through exposure to sequences of auditory events. Here, the model is configured to learn online throughout the experiment, starting with a null model, given the same experience as a typical infant in the study. The model's output at each tone position within a stimulus is a conditional (or posterior) probability distribution predicting the frequency of the next tone, given the preceding context. This distribution accumulates the model's experience during the experiment. Using the posterior distribution, the model estimates the predictability of each possible continuation tone, including the tone that follows. The model's output is formalized using the information-theoretic concept of information content (IC). IC measures the unexpectedness of a tone appearing at a particular position in the sequence given the preceding context. Here, an initially empty model is trained incrementally on a stimulus set identical to those presented to participants. Because the model learns dynamically from the entire stimulus set of short tone sequences, the model is not sensitive to alphabet size within individual stimuli. Modeling data presented here are computed by training the model on ten random and ten regular stimuli for each participant, selected randomly from the set used for the experiment. Since the IDyOM model uses a melody representation based on the Western chromatic scale (consisting of 12 semitones), whereas our stimuli used tones of a 26-tone logarithmically spaced scale in the range of 198–3563 Hz (see point 2.2 for details), we converted stimuli to IDyOM's representation by mapping the alphabet of each stimulus to an identically sized segment of the C major scale from C1 upwards.

The model can be used to determine the earliest time at which an ideal Bayesian observer can detect the repetition of the tonal pattern. This time point is quantified by identifying the tone position at which the model output (IC) differs between the REG and RAND stimuli. Therefore, the computed IC time series for the RAND and REG stimuli were entered into the same temporal cluster-based permutation test used for the GFP responses to RAND and REG stimuli (Section 2.4.2).

3. Results

3.1. Responses evoked by RAND and REG stimuli

Results are summarized in Fig. 2. The global field potential (GFP) was significantly different for RAND than REG trials in two temporal clusters: 1048–1594 ms ($p = 0.045$) and 2580–3500 ms ($p = 0.022$) from stimulus onset. Thus, the first significant time window emerged shortly after two pattern cycles were completed in the REG condition (Fig. 2b).

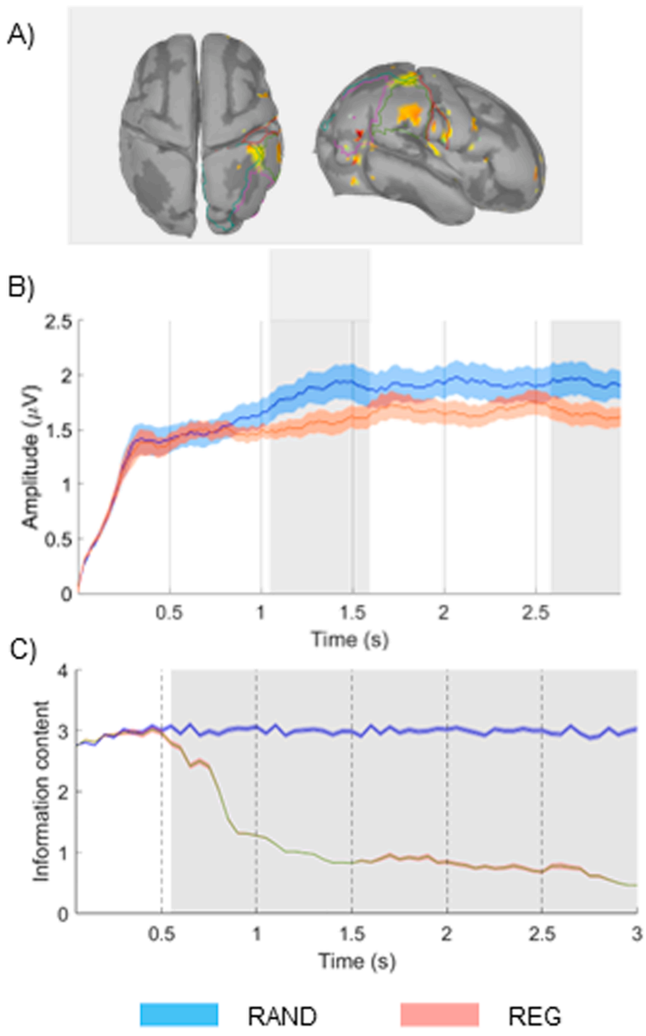

Fig. 2. a) The significant sources of electric brain activity difference (t values) between REG and RAND stimuli (between 1048 and 1594 ms) plotted on the cortical surface from the top and right views. b) Average global field power of the electric response elicited by REG (red line) or RAND (blue line) stimuli. The surrounding red/blue shaded range corresponds to the standard error. Green dashed lines show the borders between the repeated 10-tone patterns in the REG stimulus. Gray-shaded areas indicate the time windows of the significantly different temporal clusters. c) Information content (IC) time series for RAND (blue line) and REG (red line) stimuli from the IDyOM model. The gray shaded area indicates the cluster within which the two models produce significantly different IC values.

Neural generators of recurrent sound pattern recognition processes are investigated in the first-time window of significant GFP difference (between 1048 and 1594 ms). Source localization revealed that the REG vs. RAND effect (significant difference in current source density between the two types of stimuli) was dominantly generated in the right hemisphere (Fig. 2a), in the parietal cortex (Inferior parietal (IPG) and supramarginal (SMG) postcentral gyrus).

Comparing the information content time series produced by the IDyOM models trained separately on the RAND and REG stimuli of the current experiment yielded a significant temporal cluster starting at tone 11($p < 0.001$). Thus, for the ideal Bayesian observer, REG and RAND stimuli are distinguishable already during the second cycle of the repeating pattern of the regular stimulus.

3.2. *Results of the deviant evoked response analysis*

In response to REG but not RAND stimuli, deviants elicited significantly higher-amplitude event-related potentials than non-deviants (Fig. 2). The significantly different temporal cluster was between 44 and 358 ms from the onset of the deviant ($p = 0.024$). We calculated the difference wave (deviant-standard) separately for RAND and REG conditions for the fronto-central region of interest to test the interactional effect of deviance detection and sequence pattern type. Individual peak detection was then performed between 44 and 358 ms time-locked to the onset of the deviant tone. The MMN response was entered into a single-sample t-test against 0 which was significant only in the REG condition (for REG condition, $t = 2.075$, df $= 31$, $p = 0.046$; for RAND condition, $t = 0.009$, df $= 31$, $p = 0.992$). A dependent sample t-test to compare MMN response in REG relative to RAND, which showed no significant difference between conditions ($t = 1.28$, $p = 0.210$) (Fig. 3).

4. Discussion

The present study addressed whether the newborn auditory system can automatically detect the emergence of a repeating sequential sound pattern based on statistical learning mechanisms. Repeating (thus predictable) tone patterns (REG) was associated with a slow neural response that differed from that elicited by random-tone stimuli (RAND). Further, deviant tones only elicited a significant MMR response when violating REG stimuli. Computational modeling indicated that the ideal Bayesian observer would somewhat distinguish REG and RAND sequences sooner (at ca. 500 ms) than the neonatal brain. Our results suggest that the capacity to scan the auditory environment and represent repeated patterns within it is innate, while the processes leading to detecting the structure change or get tuned during maturation, as adults

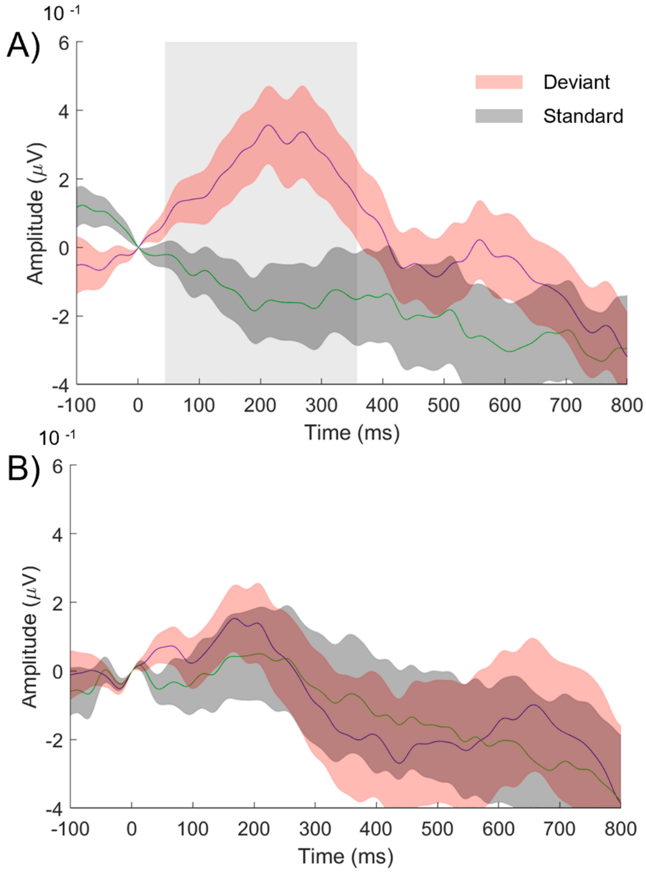

Fig. 3. Standard (red line) or Deviant (gray line) tones elicited average event-related potentials (MMR) in REG (A) and RAND (B) stimuli. The surrounding red/gray shaded range corresponds to the standard error. Gray-shaded areas indicate the time windows of the significantly different temporal clusters.

detect pattern repetition even before the end of the second cycle of the pattern (Barascud et al., 2016). Below we further discuss how the newborn auditory system tracks and represents the structure of sound sequences.

Sensitivity to repetition in tone patterns has been previously demonstrated in adults (Barascud et al., 2016; Southwell et al., 2017; Southwell and Chait, 2018). The stimulus was identical to the stimuli and procedure presented to the adults, but the newborns were asleep during the recording. There are two main differences between the electric brain activity observed in adults (Barascud et al., 2016) and in neonates to the change from a sequence of random tones to a repeating tone pattern: 1) while higher-energy responses were elicited in adults, the opposite occurred in neonates and 2) the onset of significantly diverging neural activity commenced during the second cycle of the pattern in adults, while it only started after completing the second entire cycle in neonates. These differences suggest that detecting repeating sound patterns relies on partly different neural processes in adults and neonates.

The first observation is notable because predicting coding (Friston and Kiebel, 2009) assumes that predictable input elicits less neural activity than unpredictable stimuli (no or lower error signals, no or less top-down change to models). Therefore, while interpreting the adult findings poses problems for the predictive coding framework (Heilbron and Chait, 2018), the current ones in neonates are compatible with it, as well as with previous findings of the effects of repetition in adults (Garrido et al., 2013; Wacongne et al., 2011; Winkler et al., 2009; Baldeweg, 2006; Garrido et al., 2009; Todorovic and de Lange, 2012; Kok et al., 2012). While repeating a sound may elicit lower neural responses through adaptation, comparing two stimuli composed of the same tones with only their order being different rules out this explanation. Instead, the current differential neural response is likely linked to processes detecting and encoding transitional probabilities, as were shown to operate in adults even when the auditory stimuli were not task-relevant (Mittag et al., 2016). Transitional probabilities underlie the prediction of upcoming sounds, a process observed in newborn infants (Háden et al., 2015). Thus, one potential cause of the difference between the adult and neonatal responses is that they detect or encode transitional probabilities differently. It is important to note that the vigilance state differences between awake adults and sleeping newborns may cause age-related differences in their GFP responses. It is possible that, unlike sleeping neonates, the adult subjects were attentively searching for repetitive patterns within the stimuli. To test this alternative explanation, adults would need to perform attentively demanding tasks (such as visual n-back tasks) while listening to the identical acoustic stimuli as newborns.

The other difference between adult and neonate responses to repeating a tone pattern is that the response starts earlier in adults, close to that predicted for the ideal Bayesian observer (Barascud et al., 2016; Southwell and Chait, 2018). One possible reason is that pitch resolution is much rougher in infants than adults (Fassbender, 1996; Gerken and Aslin, 2005; Novitski et al., 2007). If infants cannot distinguish all the tones from each other within the repeating pattern, it could delay the detection of its repetition. While the current study did not test patterns of different numbers of tones, it is notable that divergence between the responses to REG and RAND sequences commenced soon after the offset of the second entire cycle. This brings up another possible explanation of the difference between the adult and newborn responses to repeating tone patterns. It is possible that adults can utilize fragments of the pattern to treat the sequence as a repeating pattern, as was suggested in the computational model of Harrison and colleagues (2020) and some findings about the perception of periodic noise (Ringer et al., 2023), newborn infants rely only on detecting complete pattern repetition. This also explains the difference between the newborn brain and the ideal Bayesian observer.

Results of the source analysis indicated that in neonates, regularity detection is subserved by a right hemispheric network comprising PAC, SMG, and PoCG. This is generally consistent with previous reports on the neural generators linked with auditory sequence learning and implicit learning. SMG was reportedly involved in short-term memorization and comparison of acoustic features of the incoming auditory stimuli (Celsis et al., 1999). Barascud and colleagues identified EEG and fMRI, the PAC and IFG as the primary substrates of detecting auditory pattern repetition. However, they focused on the initial part of REG-RAND divergence. One possibility is that the contribution of the IFG to regularity detection is limited in time and restricted to the early stages of pattern detection, or it only gets involved during development.

Responses to deviants further distinguished REG and RAND sequences, as only deviant stimuli embedded in REG sequences elicited a response that significantly differed from that elicited by standard ones. Because deviance came from the tone frequency being outside the range of the other tones in the stimuli, we hypothesized that deviants would elicit the MMR response in both conditions, with the REG deviants eliciting larger MMR, because they also violated the regular repetition of the tone pattern (being presented only after the newborn brain detected repetition). Studies in adults showed that MMN is elicited by deviants falling outside the frequency range within which standard tones vary (Daikhin and Ahissar, 2012; Winkler, 1990). The fact that no MMR was elicited in the RAND condition suggests that the current deviation from the standard range was not sufficiently distinctive or that the variation of the standard tones was too considerable for forming a clear feature standard. On the other hand, MMR elicitation in the REG condition strongly supports the conclusion that pattern repetition was encoded in the neonate's brain, and tones deviating from this pattern were detected.

The present auditory stimuli provide a basic model for learning of highly structured tone sequences, such as music and language. Although previous studies have identified different potential learning mechanisms in older infants (Gomez and Gerken, 1999; Saffran et al., 1996; Marcus et al., 1999), the earliest sensitivities remain unexplored. To address this issue, a study (Gervain et al., 2008) found in newborns increased responses to the repetition sequences in the temporal and left frontal areas for repetition-based structures (syllable sequences containing immediate repetitions (ABB; e.g., "mubaba,'' "penana'') relative to random control sequences (ABC; e.g., "mubage,'' "penaku''). This also suggests an automatic perceptual mechanism to detect repetitions in line with the present results. This common perceptual ability may serve the origin and facilitate later language development.

Funding

This work was funded by the Hungarian National Research Development and Innovation Office (ANN131305 and PD123790 to BT, and K115385 to IW); the János Bolyai Research Grant awarded to BT (BO/00237/19/2) and separately to GPH (BO/00523/21/2) and the New National Excellence Program of the Ministry for Innovation and Technology from the source of the National Research, Development and Innovation (ÚNKP-22-5-BME-319) for GPH. We want to thank Eszter Rozgonyiné Lányi for their assistance during data collection.

CRediT authorship contribution statement

Brigitta Tóth: Conceptualization, Methodology, Data curation, Writing – original draft, Writing – review & editing, Funding acquisition, Project administration. **Péter Kristóf Velősy:** Methodology, Formal analysis. **Petra Kovács:** Writing – original draft. **Gábor Peter Háden:** Conceptualization, Methodology, Writing – review & editing. **Silvia Polver:** Formal analysis. **Istvan Sziller:** Data curation, Funding acquisition, Project administration. **István Winkler:** Conceptualization, Methodology, Writing – review & editing.

Declaration of Competing Interest

The authors declare that the research was conducted without any commercial or financial relationships that could be construed as a potential conflict of interest.

Data availability

Data will be shared on OSF website. Data is available at https://osf.io/834fm/.

References

Alho, K., Sainio, K., Sajaniemi, N., Reinikainen, K., Näätänen, R., 1990. Event-related brain potential of human newborns to pitch change of an acoustic stimulus. Electroencephalogr. Clin. Neurophysiol./Evok. Potent. Sect. 77 (2), 151–155.

Anders, T.F., Emde, R.N., Parmelee, A.H., 1971. A Manual of Standardized Terminology, Techniques and Criteria for Scoring of States of Sleep and Wakefulness in Newborn Infants. UCLA Brain Information Service/BRI Publications Office, NINDS Neurological Information Network, Los Angeles, CA.

Baldeweg, T., 2006. Repetition effects to sounds: evidence for predictive coding in the auditory system. Trends Cogn. Sci. (Regul. Ed.) 10 (3), 93–94.

Barascud, N., Griffiths, T.D., McAlpine, D., Chait, M., 2014. "Change deafness" arising from inter-feature masking within a single auditory object. J. Cogn. Neurosci. 26 (3), 514–528.

Barascud, N., Pearce, M.T., Griffiths, T.D., Friston, K.J., Chait, M., 2016. Brain responses in humans reveal ideal observer-like sensitivity to complex acoustic patterns. Proc. Natl. Acad. Sci. 113 (5), E616–E625.

Basirat, A., Dehaene, S., Dehaene-Lambertz, G., 2014. A hierarchy of cortical responses to sequence violations in three-month-old infants. Cognition 132 (2), 137–150.

Bekinschtein, T.A., Dehaene, S., Rohaut, B., Tadel, F., Cohen, L., Naccache, L., 2009. Neural signature of the conscious processing of auditory regularities. Proc. Natl. Acad. Sci. 106 (5), 1672–1677.

Bendixen, A., SanMiguel, I., Schröger, E., 2012. Early electrophysiological indicators for predictive processing in audition: a review. Int. J. Psychophysiol. 83 (2), 120–131.

Bendixen, A., 2014. Predictability effects in auditory scene analysis: a review. Front. Neurosci. 8, 60.

Celsis, P., Boulanouar, K., Doyon, B., Ranjeva, J.P., Berry, I., Nespoulous, J.L., Chollet, F., 1999. Differential fMRI responses in the left posterior superior temporal gyrus and left supramarginal gyrus to habituation and change detection in syllables and tones. Neuroimage 9 (1), 135–144.

Daikhin, L., Ahissar, M., 2012. Responses to deviants are modulated by subthreshold variability of the standard. Psychophysiology 49 (1), 31–42. https://doi.org/10.1111/j.1469-8986.2011.01274.x.

Dale, A.M., Liu, A.K., Fischl, B.R., Buckner, R.L., Belliveau, J.W., Lewine, J.D., Halgren, E., 2000. Dynamic statistical parametric mapping: combining fMRI and MEG for high-resolution imaging of cortical activity. Neuron 26 (1), 55–67. https://doi.org/10.1016/S0896-6273(00)81138-1.

Delorme, A., Sejnowski, T., Makeig, S., 2007. Enhanced detection of artifacts in EEG data using higher-order statistics and independent component analysis. Neuroimage 34 (4), 1443–1449.

Desikan, R.S., Ségonne, F., Fischl, B., Quinn, B.T., Dickerson, B.C., Blacker, D., Buckner, R.L., Dale, A.M., Maguire, R.P., Hyman, B.T., Albert, M.S., Killiany, R.J., 2006. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. NeuroImage 31 (3), 968–980. https://doi.org/10.1016/j.neuroimage.2006.01.021.

Fassbender, C., 1996. Infants' auditory sensitivity towards acoustic parameters of speech and music. In: Delière, Irene, Sloboda, John (Eds.), Musical Beginnings: Origins and Development of Musical Competence. Oxford. https://doi.org/10.1093/acprof:oso/9780198523321.003.0003.

Fitzgerald, K., Todd, J., 2020. Making sense of mismatch negativity. Front. Psychiatry 11, 468.

Friston, K., 2005. A theory of cortical responses. Philos. Trans. R. Soc. B: Biol. Sci. 360 (1456), 815–836.

Friston, K., Kiebel, S., 2009. Predictive coding under the free-energy principle. Philos. Trans. R. Soc. B: Biol. Sci. 364 (1521), 1211–1221.

Garrido, M.I., Kilner, J.M., Kiebel, S.J., Friston, K.J., 2009. Dynamic causal modeling of the response to frequency deviants. J. Neurophysiol. 101 (5), 2620–2631.

Garrido, M.I., Sahani, M., Dolan, R.J., 2013. Outlier responses reflect sensitivity to statistical structure in the human brain. PLoS Comput. Biol. 9 (3), e1002999.

Gerken, L., Aslin, R.N., 2005. Thirty years of research on infant speech perception: the legacy of Peter W. Jusczyk. Lang. Learn. Dev. 1 (1), 5–21.

Gervain, J., Macagno, F., Cogoi, S., Peña, M., Mehler, J., 2008. The neonate brain detects speech structure. Proc. Natl. Acad. Sci. 105 (37), 14222–14227.

Gomez, R.L., Gerken, L.A., 1999. Artificial grammar learning by 1-year-olds leads to specific and abstract knowledge. Cognition 70, 109–135.

Gramfort, A., Papadopoulo, T., Olivi, E., Clerc, M., 2011. Forward field computation with OpenMEEG. Comput. Intell. Neurosci. 2011, 923703.

Háden, G.P., Németh, R., Török, M., Winkler, I., 2015. Predictive processing of pitch trends in newborn infants. Brain Res. 1626, 14–20. https://doi.org/10.1016/j.brainres.2015.02.048.

Harrison, P.M., Bianco, R., Chait, M., Pearce, M.T., 2020. PPM-Decay: a computational model of auditory prediction with memory decay. PLoS Comput. Biol. 16 (11), e1008304 https://doi.org/10.1371/journal.pcbi.1008304.

Heilbron, M., Chait, M., 2018. Great expectations: is there evidence for predictive coding in auditory cortex? Neuroscience 389, 54–73.

Jacobsen, T., Schröger, E., Horenkamp, T., Winkler, I., 2003. Mismatch negativity to pitch change: varied stimulus proportions in controlling effects of neural refractoriness on human auditory event-related brain potentials. Neurosci. Lett. 344 (2), 79–82.

Kok, P., Rahnev, D., Jehee, J.F., Lau, H.C., De Lange, F.P., 2012. Attention reverses the effect of prediction in silencing sensory signals. Cereb. Cortex 22 (9), 2197–2206.

Kushnerenko, E.V., Van den Bergh, B.R.H., Winkler, I., 2013. Separating acoustic deviance from novelty during the first year of life: a review of event-related potential evidence. Front. Psychol. 4, 595. https://doi.org/10.3389/fpsyg.2013.00595.

Marcus, G., Vijayan, S., Bandi Rao, S., Vishton, P.M., 1999. Rule-learning in seven-month-old infants. Science 283, 77–80.

Maris, E., Oostenveld, R., 2007. Nonparametric statistical testing of EEG-and MEG-data. J. Neurosci. Methods 164 (1), 177–190. https://doi.org/10.1016/j.jneumeth.2007.03.024.

McDermott, J.H., Schemitsch, M., Simoncelli, E.P., 2013. Summary statistics in auditory perception. Nat. Neurosci. 16 (4), 493–498.

Mittag, M., Takegata, R., Winkler, I., 2016. Transitional probabilities are prioritized over stimulus/pattern probabilities in auditory deviance detection: memory basis for predictive sound processing. J. Neurosci. 36 (37), 9572–9579. https://doi.org/10.1523/JNEUROSCI.1041-16.2016.

Näätänen, R., Gaillard, A.W., Mäntysalo, S., 1978. Early selective-attention effect on evoked potential reinterpreted. Acta Psychol. (Amst) 42 (4), 313–329.

Novitski, N., Huotilainen, M., Tervaniemi, M., Näätänen, R., Fellman, V., 2007. Neonatal frequency discrimination in 250–4000-Hz range: electrophysiological evidence. Clin. Neurophysiol. 118 (2), 412–419. https://doi.org/10.1016/j.clinph.2006.10.008.

Paavilainen, P., Arajärvi, P., Takegata, R., 2007. Preattentive detection of nonsalient contingencies between auditory features. Neuroreport 18 (2), 159–163.

Paavilainen, P., 2013. The mismatch-negativity (MMN) component of the auditory event-related potential to violations of abstract regularities: a review. Int. J. Psychophysiol. 88 (2), 109–123.

Pearce, M.T., 2005. The Construction and Evaluation of Statistical Models of Melodic Structure in Music Perception and Composition. Doctoral dissertation, City University London.

Pearce, M.T., Ruiz, M.H., Kapasi, S., Wiggins, G.A., Bhattacharya, J., 2010. Unsupervised statistical learning underpins computational, behavioural, and neural manifestations of musical expectation. Neuroimage 50 (1), 302–313.

Pearce, M.T., Wiggins, G.A., 2006. Expectation in melody: the influence of context and learning. Music Percept. 23 (5), 377–405.

Ringer, H., Schröger, E., Grimm, S., 2023. Within-and between-subject consistency of perceptual segmentation in periodic noise: a combined behavioral tapping and EEG study. Psychophysiology 60 (2), e14174. https://doi.org/10.1111/psyp.14174. Epub 2022 Sep 15. PMID: 36106761.

Rubin, J., Ulanovsky, N., Nelken, I., Tishby, N., 2016. The representation of prediction error in auditory cortex. PLoS Comput. Biol. 12 (8), e1005058.

Saffran, J.R., Johnson, E.K., Aslin, R.N., Newport, E.L., 1999. Statistical learning of tone sequences by human infants and adults. Cognition 70 (1), 27–52.

Schröger, E., Bendixen, A., Denham, S.L., Mill, R.W., Bőhm, T.M., Winkler, I., 2014. Predictive regularity representations in violation detection and auditory stream segregation: from conceptual to computational models. Brain Topogr. 27, 565–577.

Southwell, R., Chait, M., 2018. Enhanced deviant responses in patterned relative to random sound sequences. Cortex: J. Devot. Study Nerv. Syst. Behav. 109, 92–103. https://doi.org/10.1016/j.cortex.2018.08.032.

Southwell, R., Baumann, A., Gal, C., Barascud, N., Friston, K., Chait, M., 2017. Is predictability salient? A study of attentional capture by auditory patterns. Philos. Trans. R. Soc. B: Biol. Sci. 372 (1714), 20160105.

Sussman, E., Steinschneider, M., Gumenyuk, V., Grushko, J., Lawson, K., 2008. The maturation of human evoked brain potentials to sounds presented at different stimulus rates. Hear. Res. 236 (1–2), 61–79.

Tadel, F., Baillet, S., Mosher, J.C., Pantazis, D., Leahy, R.M., 2011. Brainstorm: a user-friendly application for MEG/EEG analysis. Comput. Intell. Neurosci. 2011, 1–13.

Todorovic, A., de Lange, F.P., 2012. Repetition suppression and expectation suppression are dissociable in time in early auditory evoked fields. J. Neurosci. 32 (39), 13389–13395.

Virtala, P., Huotilainen, M., Partanen, E., Fellman, V., Tervaniemi, M., 2013. Newborn infants' auditory system is sensitive to Western music chord categories. Front. Psychol. 4, 492. https://doi.org/10.3389/fpsyg.2013.00492.

Wacongne, C., Labyt, E., Van Wassenhove, V., Bekinschtein, T., Naccache, L., Dehaene, S., 2011. Evidence for a hierarchy of predictions and prediction errors in human cortex. Proc. Natl. Acad. Sci. 108 (51), 20754–20759.

Winkler, I., Denham, S.L., Nelken, I., 2009. Modeling the auditory scene: predictive regularity representations and perceptual objects. Trends Cogn. Sci. (Regul. Ed.) 13 (12), 532–540.

Winkler, I., Paavilainen, P., Alho, K., Reinikainen, K., Sams, M., Naatanen, R., 1990. The effect of small variation of the frequent auditory stimulus on the event-related brain potential to the infrequent stimulus. Psychophysiology 27 (2), 228–235. https://doi.org/10.1111/j.1469-8986.1990.tb00374.x.

Winkler, I., Schröger, E., 2015. Auditory perceptual objects as generative models: setting the stage for communication by sound. Brain Lang. 148, 1–22.