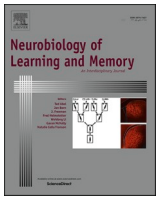
Neurobiology of Learning and Memory 215 (2024) 107987

Contents lists available at ScienceDirect

Neurobiology of Learning and Memory

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



Long-term memory formation for voices during sleep in three-month-old infants

Lisa Bastian [a,b,1], Eva-Maria Kurz [a,c,1], Tim Näher [d], Katharina Zinke [a], Manuela Friedrich [e,f,2], Jan Born [a,g,h,*,2]

[a] Institute of Medical Psychology and Behavioural Neurobiology, University of Tübingen, 72076 Tübingen, DE, Germany
[b] Max Planck School of Cognition, Max Planck Institute for Human Cognitive and Brain Sciences, 04103 Leipzig, DE, Germany
[c] Department of Child and Adolescent Psychiatry, Psychosomatics and Psychotherapy, University Hospital of Psychiatry and Psychotherapy, University of Tübingen, 72076 Tübingen, DE, Germany
[d] Ernst Strüngmann Institute (ESI) for Neuroscience in Cooperation with Max Planck Society, 60528 Frankfurt, DE, Germany
[e] Department of Neuropsychology, Max Planck Institute for Human Cognitive and Brain Sciences, 04103 Leipzig, DE, Germany
[f] Department of Psychology, Humboldt-University, 12489 Berlin, DE, Germany
[g] Center for Integrative Neuroscience, University of Tübingen, 72076 Tübingen, Germany
[h] German Center for Mental Health (DZPG), site Tübingen, Germany

ARTICLE INFO

Keywords:
Event related potential (ERP)
Early infancy
Long-term memory
Mismatch response (MMR)
Sleep

ABSTRACT

The ability to form long-term memories begins in early infancy. However, little is known about the specific mechanisms that guide memory formation during this developmental stage. We demonstrate the emergence of a long-term memory for a novel voice in three-month-old infants using the EEG mismatch response (MMR) to the word "baby". In an oddball-paradigm, a frequent standard, and two rare deviant voices (novel and mother) were presented before (baseline), and after (test) familiarizing the infants with the novel voice and a subsequent nap. Only the mother deviant but not the novel deviant elicited a late frontal MMR (~850 ms) at baseline, possibly reflecting a long-term memory representation for the mother's voice. Yet, MMRs to the novel and mother deviant significantly increased in similarity after voice familiarization and sleep. Moreover, both MMRs showed an additional early (~250 ms) frontal negative component that is potentially related to deviance processing in short-term memory. Enhanced spindle activity during the nap predicted an increase in late MMR amplitude to the novel deviant and increased MMR similarity between novel and mother deviant. Our findings indicate that the late positive MMR in infants might reflect emergent long-term memory that benefits from sleep spindles.

1. Introduction

The infant brain is highly plastic and capable of forming long-term memories (DeCasper & Fifer, 1980). Infants as young as three months can demonstrate associative learning and retention after receiving reminders across different delays (Alberini & Travaglia, 2017; Hayne et al., 2000). However, the mechanisms by which the infant brain forms such long-term representations and the role of sleep in this process (Huber & Born, 2014) is still poorly understood. Neural correlates of long-term memory representations have been investigated through event-related potentials (ERPs) in EEG recordings with so-called oddball paradigms. These paradigms involve the repeated and intermixed presentation of frequent "standard" stimuli and rare "deviant" stimuli. The difference between the ERPs to the rarely presented deviant and frequently presented standard stimuli is typically calculated to determine the Mismatch Response (MMR).

The MMR is modulated by the familiarity of the deviant stimulus and has thus been implicated in long-term memory-based processing. For example, newborns express an increased MMR to familiarized stimuli during sleep and wake states (Beauchemin et al., 2011; Cheour et al., 2000; Cheour et al., 2002; Partanen et al., 2013). At three months, infants show a late positive MMR that was distinctly increased to words spoken by the mother compared to a stranger (Zinke et al., 2018). These findings highlight the involvement of long-term memory-based

* Corresponding author at: Institute of Medical Psychology and Behavioral Neurobiology, University of Tübingen, Germany.
E-mail address: jan.born@uni-tuebingen.de (J. Born).
[1] Shared first authorship.
[2] Shared senior authorship.

https://doi.org/10.1016/j.nlm.2024.107987
Received 12 November 2023; Received in revised form 20 August 2024; Accepted 11 September 2024
Available online 14 September 2024

processes shortly after birth and increased processing of socially relevant stimuli at this early age. While previous studies largely focused on the MMRs associated with established memory representations, it remains unclear whether the MMR can also map the process of forming new long-term memory.

It is well known that sleep benefits the formation of long-term memories (Brodt et al., 2023; Diekelmann & Born, 2010). Sleep spindles have been identified as mechanisms that support memory consolidation and underlying synaptic plasticity in cortical circuits (Bastian et al., 2022; Gais et al., 2002; Klinzing et al., 2019; Latchoumane et al., 2017; Timofeev et al., 2002). Spindles are also suspected of supporting memory consolidation and plasticity during infancy (Huber & Born, 2014). Although rare, spindle activity is already present in infants younger than one month and increases until around four months, reaching a rate of 2–3 events per minute with a frequency of around 13 Hz over frontopolar and central regions (Kwon et al., 2023). Numerous studies have shown that sleep spindles are associated with indicators of long-term memory processing as early as six months of age (Friedrich et al., 2022; Friedrich et al., 2019, 2020; Friedrich et al., 2015; Friedrich et al., 2017; Kurdziel et al., 2013). However, it is still unclear whether the mechanisms by which sleep spindles aid memory are consistent between infants and adults (Mason et al., 2021; Seehagen, 2019).

In this study, we used the MMR to explore how three-month-old infants form long-term memories of an initially unknown voice. Our research builds on Zinke et al. (2018), who demonstrated infants' differential processing of their mother's voice versus a novel voice. Both voices were infrequently presented deviant stimuli in a mismatch paradigm. Specifically, Zinke et al. (2018) found that the novel voice elicited a negative deflection in the MMR in an early latency range (200–300 ms post-stimulus) indicating novelty processing in short-term memory. In contrast, the mother's voice elicited an increased MMR in a later latency range (300–400 ms), suggesting long-term memory processing due to the high familiarity of the voice to the infant. We seek to extend the findings by Zinke et al. (2018) by investigating the process of long-term memory formation. In addition to the initial voice mismatch task, we incorporated a familiarization phase for the novel voice, subsequent sleep, and a second presentation of the voice mismatch task in this study (cf. Fig. 1a for the experimental design). We hypothesized that (1) the late MMR to the mother's voice would remain consistent from the initial to the subsequent presentation of the voice mismatch task, indicating stable long-term memory processing of the mother's voice. Additionally, we expected (2) the formation of a long-term memory representation for the novel voice after voice familiarization and sleep. To test this, we directly compared the MMRs to the mother's and novel voices at baseline and test to assess the change in similarity of the two MMRs following familiarization and sleep. Assuming that sleep particularly facilitates long-term memory formation, we also hypothesized (3) that neural correlates of long-term memory in the MMR would correlate with spindle activity during sleep.

In line with our three hypotheses, we show that infants develop an MMR to the novel voice after a period of voice familiarization and subsequent sleep. The MMR exhibits a robust late positive shift and an early transient negativity, strikingly similar to the MMR in response to the mother's voice. Crucially, the emergence of a late positive shift in the MMR was positively associated with the number of sleep spindles during the nap that followed the voice familiarization. Additionally, the enhanced similarity between the MMRs to the novel voice and mother's voice after familiarization and sleep was positively associated with spindle amplitude.

2. Materials and methods

2.1. Participants

The initial sample consisted of 35 infants aged between 10 and 18 weeks, of which 31 participants had already been included in the previous publication by Zinke et al. (2018), partially comprising the data of our baseline session. However, we excluded a total of fifteen infants from our data analysis because of poor signal quality ($n = 7$), falling asleep during familiarization ($n = 1$), falling asleep during the second presentation of the voice mismatch task ($n = 1$), no toleration of the EEG cap ($n = 2$), insufficient sleep (<10 min) or not falling asleep within 30 min during the post-familiarization sleep period ($n = 4$). Twenty infants (12 females, mean age 99.8 ± 3.5 days; range: 74 to 129 days) remained in the final sample. All infants in the final sample slept at least 20 min during the break, i.e. after voice familiarization, and adhered to the study protocol. All infants were born singleton at full-term (mean gestational age: 39.85 ± 0.2 weeks) with normal neonatal outcome (birth weight >2500 g, mean birth weight: 3530 ± 98 g, mean birth height 52 ± 0.4 cm), were healthy according to parental report, and had no severe complications during pregnancy or delivery. The whole sample had an Apgar score above 9 at 10 min after birth (median 9/9/10 for the 1/5/10 min Apgar score). Parents of infants were recruited via email advertisements across the university's mailing system, flyers, and through mothers who had already participated in another study during pregnancy with their child. Participating families received monetary compensation for their time and effort. The study was approved by the ethics committee of the Medical Faculty of the University of Tübingen.

2.2. Voice mismatch task and voice familiarization

Our voice mismatch task is a variant of the classical oddball paradigm and was adapted from previous studies that investigated the discrimination of familiar and unfamiliar voices in newborns (Beauchemin et al., 2011; Mai et al., 2012). Thus, the task was not developed for investigating discrimination ability of different words, but different voices saying the same word. The recordings of the word "baby" (400 ms, ISI = 600 ms) were repeatedly presented and pronounced by three different female speakers: an unfamiliar one speaking the frequently presented "standard" voice (85 % of the trials, $n = 510$), the infant's own mother pronouncing the infrequently presented "mother deviant" voice (7.5 % of the trials, $n = 45$), and a second unfamiliar voice pronouncing the rarely presented "novel deviant" (7.5 % of the trials, $n = 45$). For each infant, the standard and novel deviant voices were chosen from a pool of four female voice recordings with the allocation of voices being balanced across participants. Stimuli were presented in a pseudorandomized order using the software Presentation (Neurobehavioral Systems, Berkeley, USA). Stimulus presentation lasted for a total of 10 min (600 trials total) and was played to the awake infant. The stimuli were presented through loudspeakers at a constant volume of ~75 dB SPL to avoid differences in ERPs due to differences in intensity of the stimuli. For recording the task stimuli, the infants' mothers and the four other female speakers were instructed to pronounce the word "baby" with a German pronunciation as naturally as possible while avoiding any emotional connotations. The voices were recorded with a portable USB Condensor Microphone (Go Mic by Samson Technologies ®) and a custom-made pop filter using the software Audacity 2.0.5 for recording and post-processing. Minimal processing was applied to produce stimuli of comparable length and loudness using noise removal, amplifying, cutting recordings and minimally changing tempo where necessary.

To familiarize the infants with the voice of the rarely presented novel deviant, they listened to a recording of this novel deviant reading a children's story for 10 min after the baseline session. These recordings were made before the study began and were chosen based on the voice used as the novel deviant in the mismatch voice paradigm. While the mother and novel deviant voices were the same during test and baseline, a new standard stimulus was used to avoid habituation effects to this voice across sessions.

2.3. Procedures

Briefly, the experimental procedure consisted of a baseline session of the voice mismatch task, followed by voice familiarization with the novel voice, then a period of sleep, and finally a test session of the voice mismatch task (cf. Fig. 1a). Note that we did not introduce a wake control group due to ethical concerns regarding the administration of a sleep deprivation in these young infants. Experimental sessions were scheduled at a time when the infant was expected to be in a calm but alert state. After arrival at the lab, the infant adapted to the new environment. Meanwhile, the mother answered screening questionnaires we recorded her voice to create the individual mother deviant stimulus for the mismatch paradigm.

We administered EEG, electrooculography (EOG), electromyography (EMG) and electrocardiography (ECG). After setting up the EEG recordings the alert and comfortable infant was positioned on a diaper changing unit on its back with the head between two loudspeakers (distance of approx. 45 cm each). The mother stood in front of the unit and interacted with the infant (e.g., presenting hand puppets, blowing bubbles, changing facial expressions, etc.) to keep the infants calm and alert according to guidelines (Hoehl & Wahl, 2012). They were instructed not to talk to their child or make any kind of noises during the voice paradigm. The mothers did not wear noise canceling headphones or earplugs for them to appropriately react to any sounds their child might make and to keep the situation as natural as possible for the infant. If the infant was uncomfortable in this position, the mother held the infant in her arm during the recordings, with the position of loudspeakers adjusted accordingly. Most infants were fed right before the baseline session ($n = 18$) and test session ($n = 13$).

The first presentation of the mismatch voice paradigm, i.e., the baseline session (cf. Fig. 1a) started. For one infant in the baseline session and three infants in the test session, the presentation of the task stimuli was paused due to fussiness or changes in alertness. The baseline session was followed by the 10 min familiarization period during which the infant was familiarized with the voice of the novel deviant by listening to the prerecorded children's story. The infants were either lying on the diaper changing unit or held by their mother. During this time, electrodes were checked for the (mobile) polysomnographic recordings which started after the familiarization. In the following sleep interval (~60 min), the parent moved around freely with the infant in the lab or took a walk with a stroller so the infant could fall asleep. The test session with the second presentation of the voice mismatch paradigm started when the infant work up spontaneously and had been awake for at least 15 min to allow for sleep inertia to fade away. Additional ratings of the infants' level of sleepiness on a 10-point scale (from 1, "very awake", to 10, "asleep") were obtained from the mothers before and after the baseline session and before the test session.

2.4. Data acquisition

EEG was recorded using soft Ag/Cl electrodes attached to an infant-suitable cap (EASYCAP GmbH, Herrsching, Germany) at electrode

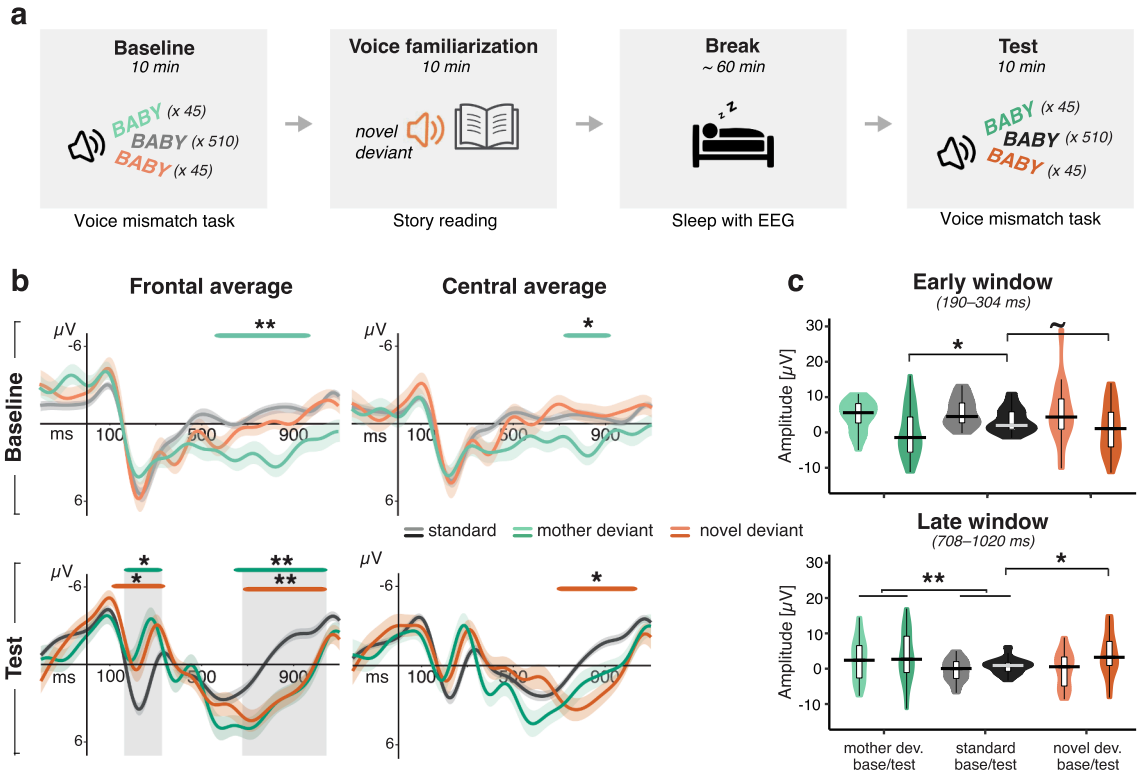


Fig. 1. Experimental design and ERP responses at baseline and test. a, Experimental procedure: Infants ($n = 20$) were tested on an auditory voice mismatch task twice, at an initial baseline and a final test session. On the voice mismatch task, the infants listened for 10 min to 3 types of voice stimuli presented in random order, i.e., the word "baby" was spoken by the mother voice (mother deviant; $n = 45$), by a novel voice (novel deviant; $n = 45$) and by a standard voice (standard; $n = 510$). The baseline session was followed by a 10 min familiarization period during which the infant listened to a story read by the novel deviant voice, and then had the opportunity to nap (for ~ 60 min). Around 15 min after awakening the test session took place. Different female voices were used as standard stimuli during baseline and test session. **b,** Grand-average ERPs from frontal (including F3, Fz, F4) and central (C3, Cz, C4) channels at baseline and test. Colored lines indicate the time windows in which the ERP response to the mother deviant voice and novel deviant voice, respectively, differed from the ERP to the standard voice (cluster-based permutation test, cluster $\alpha < 0.05$). Grey shadings indicate time windows of interest used in the analysis in c). **c,** Results of a linear mixed-effects model for the early (190–304 ms) and late (708–1020 ms) ERP time window with channel (frontal/ central), stimulus (mother deviant/ novel deviant/ standard) and session (baseline/ test) as fixed factors and participant x session as random factor. (* $p \leq 0.05$; ** $p < 0.01$). Boxplots indicate group means (horizontal bars) and the inter-quartile ranges. Results in c) are displayed collapsed over frontal and central channels.

positions F3, Fz, F4, FCz, C3, Cz, C4, Pz, and mastoids, with reference to M2, and Fp2 as ground. Electrode impedances were mostly below 10 kΩ, but always below 20 kΩ, which is considered acceptable for infant ERPs (Hoehl & Wahl, 2012). EOG recordings included one electrode below the left eye and one at the Fp1 position. The EMG was recorded from two electrodes on the chin. Signals were digitized at a rate of 500 Hz using a standard amplifier (BrainAmps, Brain Products GmbH, Gilching, Germany), and with a portable amplifier and recording system (SOMNOscreen™ plus Neuro+, SOMNOmedics GmbH, sampling rate: 256 Hz) for polysomnographic recordings. EEG recordings from Pz had to be excluded from all analyses as this channel contained too many artifacts in most participants (since most infants were lying on the back of their heads).

2.5. Data processing

2.5.1. EEG processing for ERP analyses

EEG data of the mismatch voice paradigm was preprocessed using EEP V3.2.1. (available as EEProbe, MPI for Human Cognitive and Brain Sciences, Leipzig, Germany) software. The EEG was offline re-referenced to linked mastoid electrodes and the signal was band-pass filtered between 0.44–20 Hz using a digital zero-phase Butterworth filter (−3 dB cut-off at 0.54 and 19.90 Hz). The strong DC-suppression of this filter (−80 dB) allowed the calculation of ERPs without baseline correction. In seven infants, one channel with longer periods of artifacts was mean-interpolated based on the two adjacent channels. The signal was then segmented into epochs of 1300 ms from 200 ms pre-stimulus to 1100 ms post-stimulus. Trials exceeding a standard deviation of 100 μV within a sliding window of 300 ms at any channel were rejected. ERPs were averaged time-locked to the onset of the word stimulus. A minimum of 15 artifact-free trials per condition was required for the inclusion of an individual in further analyses. For the baseline session, the mean (±SD) number of trials was 35.75 (±5.44) for the rarely presented mother deviant voice, and 35.45 (±5.20) for the rarely presented novel deviant voice, and 401.95 (±55.64) for the frequently presented standard stimulus. The respective number for the test session were 31.60 (±7.29), 31.65 (±8.31), and 349.90 (±82.07). Trial numbers were higher in the baseline than the test session ($F(1,19) = 5.00$, $p = 0.037$). Within sessions, trial numbers did not differ between the two deviant conditions (baseline: $t(19) = 0.413$, $p = 0.684$, test: $t(19) = -0.63$, $p = 0.951$). For illustrative purposes, ERPs shown in the figures were additionally filtered using a 7 Hz low-pass filter (zero-phase, Butterworth, order 4).

2.5.2. Sleep EEG processing

Offline processing of sleep EEG data was done using Brain Vision Analyzer 2.0 Software (Brain Products GmbH, Gilching, Germany), and consisted of re-referencing of EEG recordings to the linked mastoid electrodes, and filtering of EEG and EOG channels (band-pass,0.3–35 Hz) and of the EMG (band-pass,10–100 Hz). To remove line noise, a 50 Hz notch filter was applied to all channels. Sleep stages were determined for subsequent 30-sec epochs, using standard criteria recommended for children by the American Association of Sleep Medicine (AASM, Berry et al., 2017). Beforehand, each recording was screened for the necessary sleep characteristics to differentiate NonREM sleep into N1 to N3 sleep. Recordings were scored by two scorers with an interrater agreement of 89 %. For each infant, the total sleep time (TST) of the nap, time spent in different sleep stages (in minutes and in percentage of TST), and wake after sleep onset (WASO) were determined. Channels and epochs with artifacts were manually marked during the scoring process and excluded from further sleep-EEG analyses.

For spindle detection, we calculated for each participant the average power spectra during NonREM sleep (i.e., N2 and N3 sleep) for frontal and central channels. The power spectra were computed using MATLAB's *pwelch* function, with a window size of 5 sec (Hamming taper) and a 50 % overlap. Frontal and central power peaks in the spindle frequency band were separately detected and modelled as Gaussian deviations from the aperiodic 1/f components (Donoghue et al., 2020). One infant did not exhibit a clear spindle frequency peak and was excluded from further analyses of sleep spindles. If a participant had a slow and a fast spindle peak, the frequency peak with the higher power was used for detection (cf. Table S6). Nine of the 20 participants exhibited two peaks in the frontal channel, seven participants showed two peaks in the central channel. None of the identified peaks was below 12 Hz, a typical characteristic of spindles during infancy (e.g., Kwon et al., 2023). Sleep spindles were detected during artifact-free NonREM sleep (N2 und N3 sleep) using SleepTrip (RRID:SCR_017318) in MATLAB 2022a. For detection, the signal of each channel was bandpass filtered (zero-phase, Butterworth, order 4) ± 1.5 Hz around the individually determined frontal and central spindle frequency peak, respectively. The root mean square (RMS) of the filtered signal was then calculated at each sample point using a sliding 200-ms window, which was further smoothed by a 200 ms moving average window. A spindle was detected, whenever the smoothed RMS signal exceeded a threshold of 1.5 times the SD of the filtered signal in the respective channel for a duration of 0.5 to 3 sec. For each participant and channel, spindle density, duration and amplitude (max. trough to max. peak potential, in μV) were determined and then averaged across frontal (F3, Fz, F4) and central (C3, Cz, C4) channels. Because previous studies showed a difference in occurrence (Clawson et al., 2016; Kwon et al., 2023) and relevance for memory formation (Friedrich et al., 2017) between frontal and central spindles in three-months-old infants, we analyzed frontal and central spindles separately.

2.6. Data analysis

Further analyses were conducted using Fieldtrip (Oostenveld et al., 2011) in MATLAB R2022a *(Mathworks Inc*., Sherbom, Massachusetts), R version 4.1.2 (R Core Team, 2021) in RStudio V2021.09.1 (*RStudio Inc*., Boston, Massachusetts), and the pyRiemann package for covariance-based clustering (https://doi.org/10.5281/zenodo.593816) in Python 3.10.6. A p-value of < 0.05 was considered significant and results are reported as means (±SEM). Akaike Information Criterion (Akaike, 1973), Bayesian Information Criterion (Schwarz, 1978) and log-likelihood were used to assess the model fit of the linear models using a stepwise backward elimination procedure. When appropriate, p-values were corrected for multiple comparisons using Bonferroni correction.

2.6.1. ERP analysis

First, we computed averages of the individual ERPs per condition and channel, i.e., standard voice, mother deviant and novel deviant voice in the baseline and test session. Next, time windows were determined in which the ERP responses to the mother deviant and to the novel deviant voice, differed significantly from that to the standard voice in frontal and central channels. For this, dependent samples t-tests were computed for each timepoint in the entire 0–1100 ms post-stimulus interval. We controlled for multiple comparisons through cluster-corrected permutation tests with 5000 permutations (Maris & Oostenveld, 2007). To reduce noise-related variance, the group-level ERPs were further averaged over frontal and central channels. On the frontal and central average ERP waveforms we repeated the time window analysis to determine response differences between each deviant and the standard voice. Two ERP time windows were identified (early: 190–304 ms; late: 704–1010 ms) related to previous MMR findings on early-latency saliency detection in neonates (Dehaene-Lambertz & Pena, 2001; Kushnerenko et al., 2007; Leppanen et al., 2004) and late-latency memory representations for word stimuli in infants (Friedrich et al., 2009; for a review, see Kushnerenko et al., 2013). For each time window, average ERP amplitudes were analyzed as dependent variables using a linear mixed-effects model. Model predictors included *session* (baseline, test) and *stimulus* (standard, mother deviant, novel deviant) and *channel* (frontal, central) as fixed effects. For each participant, random intercepts and slopes were included in the model.

2.6.2. MMR similarity analysis

Building on the ERP analysis, MMRs were computed as a measure for long-term memory-based processing of deviant stimuli. MMRs were calculated for each participant as the difference ERP waveforms between the respective deviant voice (mother, novel) and the standard voice, for frontal and central channels at baseline and test sessions. This resulted in two MMRs for each session, i.e., one for the mother deviant and one for the novel deviant. The subsequent similarity analysis focused on frontal MMRs. For this, the MMRs of each participant were averaged over channels F3, Fz and F4. We performed the MMR similarity analysis over the average frontal channels, since we only found a consistent MMR for the mother as well as the novel deviant in both latency windows at the frontal sites.

A comprehensive analysis of global similarity across the entire 1100 ms post-stimulus interval for frontal MMR waveforms was executed through covariance matrix clustering. We evaluated the similarity across various conditions and sessions (including mother deviant and novel deviant, both at baseline and test) by inspecting the distances between channel-covariance matrices of the MMR waveforms. We used the PyRiemann package (dev0.5) to compute regularized covariance matrices for frontal EEG channels. To increase the number of data points for subsequent clustering, the MMRs were bootstrapped at the participant level, yielding 100 covariance matrices. Clusters of MMR similarity ($n = 2$, width: 2.5 SD) were then discerned using K-means. Given that covariance matrices are positive symmetric definite and hence reside on a no-Euclidean Riemannian manifold, we utilized the Riemannian distance between matrices as the K-means algorithm's distance metric. High-dimensional covariance matrices were then rendered in a two-dimensional Euclidean space using Laplacian Eigenmaps for visualization.

A local similarity examination of average frontal MMR waveforms was performed within two delineated time windows (early: 220–382 ms, late: 686–918 ms) of the MMR. These time windows corresponded to clusters where the MMRs of each session exhibited significant correlations. Pearson correlations were computed for each MMR waveform timepoint across participants, separately for baseline and test sessions, ensuring the control for multiple comparisons through cluster-corrected permutation testing with 5000 permutations. For each identified time window and condition, we calculated the mean MMR amplitude across timepoints. The similarity was then assessed as the pairwise Euclidean distance between the mean MMR amplitudes across all conditions. Finally, dependent samples *t*-tests were performed for each time window to assess whether the distances between mother-deviant MMR amplitude and novel-deviant MMR amplitude differed significantly between the baseline and test sessions.

2.6.3. Statistical analyses of sleep spindles

Sleep spindle measures (frequency, count, density, amplitude, and duration) were analyzed using linear mixed-effects models in the *lme4* package (Bates et al., 2015). Each model consisted of the predictors age (continuous) and channel (frontal vs central), as well as random intercepts for each participant. Satterthwaite's approximation as implemented in the *parameters* package (Lüdecke, 2020) was used to determine the significance of fixed effects. Fixed effects were modeled additively and not as interaction, due to a better model fit of the simpler model.

2.6.4. Associations between changes in MMRs and sleep parameters

First, we tested whether the baseline-to-test change in MMR amplitude to the novel deviant and the mother deviant voices was linked to spindle activity during NonREM-sleep. We computed linear models for the critical late (novel: 708–1020 ms, mother: 668–1022 ms) and early (novel: 130–332 ms, mother: 109–304 ms) latency windows for both deviants. Based on previous literature (Friedrich et al., 2022; Friedrich et al., 2019; Horvath et al., 2018) and our own sleep analysis, model predictors included spindle *frequency, amplitude, count, density,* and *duration*. Since we observed an effect of age on sleep spindle amplitude and duration (Fig. S3), we also included *age, age x spindle amplitude* and *age x spindle duration* as nuisance predictors. In a stepwise backward-elimination procedure the most parsimonious model was determined for each deviant x latency window combination. The models were computed for frontal and central spindle parameters separately based on our effect of recording site (frontal vs. central) on almost all spindle parameters. Considering evidence for the relevance specifically of frontal spindle activity for memory formation in infancy (Friedrich, Molle, Born, & Friederici, 2022), we expected the changes in MMRs over frontal channels to be most strongly correlated with frontal spindle parameters.

In addition, we tested for an association between changes in MMR similarity and NonREM-sleep parameters for the two obtained time windows from the MMR local similarity analysis (early: 220–382 ms, late: 686–918 ms). For each time window, we computed an MMR Similarity Index as the negative difference between MMR similarity (d) at baseline and test session:

$$MMRSI = -[d_{test}(MMR_{novel}, MMR_{mother}) - d_{baseline}(MMR_{novel}, MMR_{mother})],$$

where d represents the pairwise Euclidean distances between the MMRs to the novel and the mother deviant voice. The MMR Similarity Index was then included as dependent variable using a linear model defined by a stepwise backward-elimination procedure. Model predictors were again all relevant spindle parameters (*amplitude, count, density frequency* and *duration*), *age*, *age x spindle amplitude*, and *age x spindle duration*.

2.7. Control analysis

Mother-based sleepiness ratings before and after the baseline session and before the test session were compared using a linear mixed effects model to control for confounding effects of sleepiness. A one-way ANOVA for ERPs and dependent sample *t*-tests for MMRs were computed for the whole interval over frontal and central channels separately to exclude differences in the 200 ms pre-stimulus interval of ERP responses between stimulus conditions at the baseline and test sessions.

3. Results

3.1. ERP differences at baseline and at test

First, we investigated ERP differences in response to each rarely presented deviant stimulus – the mother's voice and a novel voice – and the frequently presented standard voice during the baseline session and the test session (cf. Fig. 1a for details of the experimental design). The respective changes in ERPs were analysed across the average frontal and average central electrodes (Fig. 1b; cf. Fig. S1 for comparisons at individual channels).

At baseline session, the ERP to the rarely presented mother deviant voice was positively shifted compared to the response to the frequently presented standard voice in a late latency range. This positive shift was more pronounced over frontal than central channels (frontal: 584–948 ms, cluster $p = 0.01$; central: 746–898 ms, cluster $p = 0.03$). Importantly, the ERP to the rarely presented novel deviant voice did not differ from the ERP to the standard voice at baseline. This pattern changed at test (after voice familiarization and sleep): Here, the response to the mother deviant voice as well as to the novel deviant voice (which the infant had encountered during voice familiarization) differed from the ERP to the standard voice. Specifically, the deviant MMRs were both positively shifted in the late time window over frontal channels (mother deviant voice: 668–1022 ms, cluster $p = 0.01$; novel deviant voice: 708–1020 ms, cluster $p < 0.01$). In addition, both mother- and novel-deviant ERPs showed a more transient negative potential shift in an early latency window at frontal sites, in comparison with the response to

the standard stimulus (mother's voice: 190–304 ms, cluster $p = 0.03$; deviant voice: 130–332 ms, cluster $p = 0.08$). Over central channels, these potential shifts were weaker and only the difference between novel deviant and standard voice in the late latency window reached significance (718–1016 ms, cluster $p = 0.02$). The overlapping timestamps of the significant clusters over frontal channels at the test session (early: 190–304 ms; late: 708–1020 ms) were used as latency windows of interest for further testing (Fig. 1c). Similar time windows have previously been identified to be involved in memory-based processing in infants (Beauchemin et al., 2011; Friedrich et al., 2009).

For the two latency windows of interest, we compared the standard voice, mother deviant, and novel deviant ERPs directly across sessions (baseline vs. test) and channels (frontal vs. central) using a linear-mixed effects model. For the late window, these analyses confirmed that the positive potential shift in the ERP to the mother deviant voice (relative to the standard voice) was present irrespective of the session, i.e., at baseline and at test (main effect Session: $b = 2.77$, SEM=0.91, $t(215) = 3.04$, $p < 0.01$). In contrast, the late positive shift in the ERP to the novel deviant voice emerged only at the test session but was absent at baseline (Stimulus x Session interaction: $b = -2.99$, SEM=1.29, $t(215) = -2.32$, $p = 0.02$). For the early latency window, the analyses revealed parallel Stimulus x Session interaction effects reaching significance for the negative potential shift to the mother deviant ($b = 2.74$, SEM=1.32, $t(215) = 2.08$, $p = 0.04$) and approaching significance for the novel deviant voice ($b = 2.44$, SEM=1.32, $t(215) = 1.85$, $p = 0.06$), thus, supporting that this early negative potential shift only emerged at the test session. Channel (frontal vs. central) did not affect the ERP amplitude differences (all $p > 0.05$).

3.2. Increased global and local MMR similarity at test

MMRs were defined by the difference amplitude of the ERPs to each rarely presented deviant voice and the frequently presented standard voice. Given that MMRs to both the mother's voice and the novel deviant voice were most pronounced at frontal channels, the subsequent similarity analyses concentrated on these recording sites (Fig. 2; cf. Fig. S2 for MMRs at each frontal and central channel). We hypothesized that the MMR waveform of the two deviant voice stimuli becomes more similar through voice familiarization and subsequent sleep. Therefore, we compared the two MMRs across the baseline and test sessions. Specifically, we analysed the global shape (over the entire 1100 ms post-stimulus latency window) and local shape (in the identified early and late MMR windows) of the two kinds of MMRs.

On the global level, we detected increased similarity between the frontal responses to the mother deviant at baseline, the mother deviant at test, and the novel deviant at test. These three conditions formed one similarity cluster identified with k-means clustering following channel covariance-based similarity analyses (cf. Methods for details). Within this cluster, MMRs to the mother deviant and the novel deviant at the test session overlapped most strongly. A second cluster that was orthogonal to the first in the latent embedding space, comprised only the frontal response to the novel deviant at baseline (Fig. 3a). The observed clustering highlights a change in the MMR to the novel deviant voice which is unrelated to the MMR to the mother's voice at baseline but shows distinctly increased similarity to the mother's voice at the test session after the infant had been familiarized with the novel deviant voice and subsequent sleep.

Next, we tested whether this increased global similarity between responses to the mother deviant and the novel deviant voices was driven by effects in local MMR-related latency windows. We again detected a late (686–918 ms post-stimulus) and an early (220–382 ms) time window in which the MMR waveforms to the mother deviant and the novel deviant voices were correlated (late cluster: $p = 0.01$, early cluster: $p = 0.03$). These clusters, again, were only significant at the test session (Fig. 2b) but not at the baseline session (Fig. 2a). The latency windows of significant clusters closely matched with the latency windows obtained for significant positive and negative ERP potential shifts (cf. Fig. 1b). Both MMRs to the mother deviant voice and to the novel deviant voice showed a temporally sustained positive deflection in the late latency window and a transient negative deflection in the early latency window. MMR amplitudes between the early and late time window were negatively correlated only at the test session (mother deviant: $r = -0.54$, $p = 0.02$; novel deviant: $r = -0.42 p = 0.07$), but not at the baseline session (mother deviant: $r = -0.03$, $p = 0.89$; novel deviant: $r = -0.25$, $p = 0.31$). This pattern suggests a link between the cognitive processes underlying the MMR waveforms in the two time windows at test (Fig. 2c).

To quantify the similarities of the MMRs in the two latency windows, we computed the average pairwise Euclidean distances between the MMR amplitudes to the two deviants at baseline and test at each timepoint in the early and late latency window, respectively. The analysis confirmed that for both time windows, MMRs to mother deviant and novel deviant voice were most similar at the test session (late: M=2.38, SD=2.18, early: M=1.55, SD=1.17) and least similar at the baseline session (late: M=3.63, SD=4.42; early: M=3.73, SD=3.46, Fig. 3b). The decrease in distance (i.e., an increase in similarity) between MMRs to mother and novel deviant voice stimuli from baseline to test session was significant for both time windows (late: $t(19) = 2.16$, SEM=0.47, $p = 0.04$; early: $t(19) = 2.49$, SEM=0.53, $p = 0.02$).

3.3. Sleep after voice familiarization and age-dependent changes

Participants slept on average 64 min (SEM=6.84, cf. Table S1 for sleep macroarchitecture). Results of a linear mixed-effects model

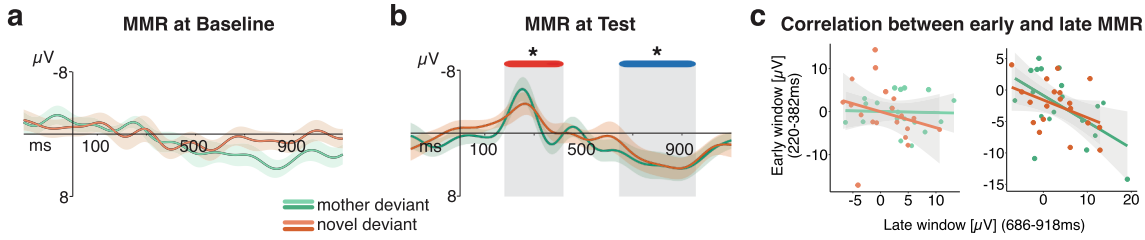


Fig. 2. Increased similarity at the test between mother and novel voice MMRs for the early and late latency windows. a, MMRs, (i.e., ERP difference waveforms between responses to each deviant and the response to the standard voice) to the mother deviant (green) and novel deviant (orange) at baseline and **b,** test sessions. Colored bars indicate time windows in which the MMR between mother deviant and novel deviant voices significantly correlated (cluster-based permutation test, cluster $\alpha < 0.05$, one-sided, positive clusters only). Grey shadings indicate time windows of interest used in the analysis in Fig. 2c and 3b. (* $p < 0.05$). **c,** Pearson correlations of average MMR amplitudes for each deviant (mother, left, novel, right) between the early (220–382 ms) and late (686–918 ms) latency windows of the MMR shown in b. *Left panel*: Results for the baseline session in which the MMR amplitudes were not correlated ($n = 20$). *Right panel*: Results for the test session, in which MMR amplitude correlations between the time windows were significant for the mother deviant, and marginally significant for the novel deviant ($n = 20$). Light green/orange colours pertain to the baseline session and dark orange/green colours to the test session. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

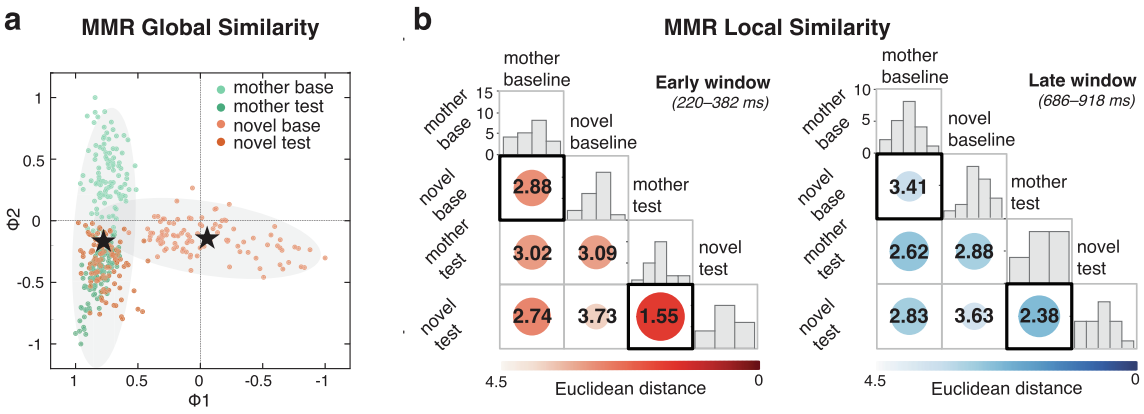


Fig. 3. Comparison of global and local MMR similarity between baseline and test sessions. a, Similarity analysis of the global MMR waveforms (i.e., the whole 1100 ms post-stimulus interval) showing MMR similarity clusters across deviant voices (mother deviant, novel deviant) and sessions (baseline, test). Clusters were obtained using K-means clustering of the bootstrapped ($n = 100$) sample covariance matrix across conditions. Asterisks indicate centroids of the clusters that each have a width of 2.5 SD. Mother deviant at baseline and test, as well as the novel deviant at test formed one similarity cluster. The other distinct cluster only comprised the novel deviant MMRs at baseline. Phis represent normalized latent variables of the 2D embedding space. **b**, Similarity analysis of the local MMR latency windows. The lower triangle of the distance matrices shows mean pairwise Euclidean distances as a measure of similarity, between MMR amplitudes across all conditions, for the identified early (220–382 ms) and late latency window (686–918 ms; cf. Fig. 1b). Histograms display the sample distribution of MMR amplitudes for each condition. Black frames indicate the MMR distances between mother and novel deviant which differed significantly between baseline and test for both latency windows ($p < 0.05$).

revealed that spindle density and amplitude were higher and spindle duration was longer at central than frontal channels (all $p < 0.02$), whereas spindle frequency was generally higher at frontal than central channels ($p = 0.027$, Table S2). There was no effect of channel ($p = 0.1$) on spindle count. Although the age range of our sample was rather narrow, we detected several age-related changes in spindle characteristics (Table S3, Fig. S3). Spindle density tended to increase with age ($p = 0.09$). For spindle amplitude and duration, the age-related increases were significant ($p < 0.001$, and $p = 0.01$, cf. Table S3 for the whole model output, Fig. S3).

3.4. Spindle activity predicts the emergence of MMRs

We used hypotheses-informed correlational analyses to assess possible contributions of sleep to MMR dynamics. Duration of the nap was positively correlated with the emergence of the late MMR to the novel deviant voice ($r = 0.52$, $p = 0.02$, uncorrected for multiple comparisons). Based on evidence for an involvement of sleep spindles in memory formation during sleep in infants (e.g., Friedrich et al., 2022; Friedrich et al., 2019), we further hypothesized that the emergence of an MMR to the novel deviant voice at the test session similar to that to the mother deviant is linked to signs of increased spindle activity during sleep following the familiarization period. To test this, we assessed the relationship between relevant measures of spindle activity (i.e., spindle count, density, frequency, duration, and amplitude) and the change from baseline to test for the MMR amplitude to the novel deviant for the critical late (708–1020 ms) and early (130–332 ms) latency window (cf. Fig. 1c). Corresponding analyses were performed for the MMR amplitude changes to the mother deviant (in the respective late 668–1022 ms and early 190–304 ms latency windows). As spindle amplitude and duration depended on the infant's age, we also included these parameters in our analysis. The analysis focused on the frontal cortical MMR where this response to the novel deviant at the test session was found to be most robust and significant. Analyses were restricted to 19 participants, because spindles were detected in all but one infant (cf. Methods section on sleep-EEG processing).

We found that frontal cortical spindle counts predicted the late (positive) MMR, i.e., the more spindles occurred over frontal regions during the post-familiarization nap the stronger the increase was in the late MMR to the novel deviant from the baseline to the test session ($t(15) = 2.86$, $p = 0.04$; Fig. 4a, Table S4). This effect of spindle count

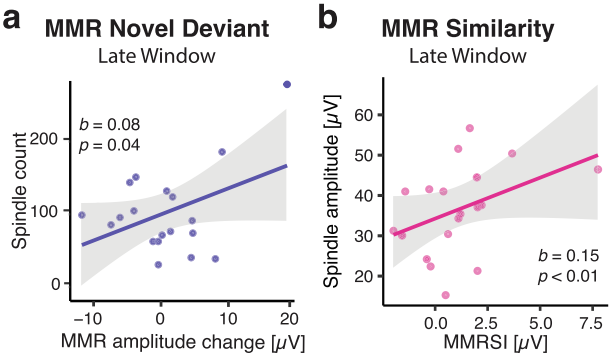


Fig. 4. Spindle activity predicts change in late positive MMR from baseline to test session. a, MMR amplitude increase in response to the novel deviant in the late latency window is positively associated with spindle count over frontal cortical channels during the post-familiarization nap ($n = 19$) **b,** Increased similarity between MMRs to mother and novel deviant voice stimuli in the late latency window is associated with higher frontal spindle amplitude during the post-familiarization nap ($n = 19$). MMR similarity is indicated by the Euclidian distance. *p*-values are corrected for multiple comparisons.

could partly be linked to the time the infants spent in NonREM sleep as in additional exploratory analyses, the late MMR to the novel deviant was found to be similarly correlated with time in NonREM sleep. However, this correlation only approached significance ($r = 0.45$, $p = 0.06$). We did not observe an effect of spindle parameters on the early (negative) MMR to the novel deviant (all $p > 0.50$). Spindle parameters over central channels were not associated with changes in MMRs to the novel deviant or mother deviant (novel deviant: late window all $p > 0.12$, early window all $p > 0.56$; mother deviant: late window all $p > 0.10$, early window all $p > 0.25$).

In addition, we investigated whether the detected similarity increase of the MMRs to the novel deviant and the mother deviant voices from baseline to test was linked to spindle activity during sleep based on an MMR Similarity Index (cf. Methods section) of the Euclidean distance between both types of MMRs for the critical late (686–918 ms) and early (220–382 ms) latency window (cf. Fig. 3b). This analysis revealed spindle amplitude in frontal cortical recordings to be a significant predictor of the increase in MMR similarity between mother deviant and

novel deviant from the baseline to the test session in the late latency window ($t(14) = 2.94$, $p = 0.01$; Fig. 4b, Table S5). For the early time window, the increase in MMR similarity across sessions was not correlated with measures of spindle activity (all $p > 0.25$).

3.5. Control measures

During the 24 h prior to the experimental session infants slept on average 13.26 h (± 0.37). Mothers' ratings of their children's sleepiness before and after the baseline session and before the test session did not significantly differ (all $p > 0.44$). Trends in the 200 ms pre-stimulus interval of ERP responses at baseline and test sessions revealed no differences between mean amplitudes across conditions for both, MMRs or ERPs (all $p > 0.05$), excluding pre-stimulus-baseline confounds in our results.

4. Discussion

We employed a mismatch voice paradigm to demonstrate how long-term memory representations of an initially unfamiliar voice emerge in three-month-old infants after a period of familiarization and subsequent sleep. ERP responses to the mother deviant voice, compared to those to the standard stimuli, revealed a distinct late (708–1020 ms post-stimulus) positive potential shift over the frontal cortex at both the baseline and the test session, likely indicating the representation of the mother's voice in long-term memory. Similarly, a late positive potential shift over frontal cortical areas was observed in the ERP response to the novel deviant voice only at the test session, following the infant's familiarization with this novel deviant and subsequent nap. Additionally, at the test session, both ERP responses to the mother deviant and novel deviant voices exhibited a transient frontocortical negative potential shift in an earlier time window (190–304 ms) compared to the ERP to the standard voices. This shift may reflect the immediate discrimination of stimulus deviance in short-term memory for both types of stimuli, which were presented rarely in comparison to the frequently-presented standard stimulus. Quantitative analysis of the MMRs (defined by the difference waveform between the ERP to the deviant voices and the ERP to the standard stimuli, for both mother and novel deviants) supported these observations. This analysis indicated that the MMR to the novel deviant at the test session after voice familiarization notably increased in similarity to that of the mother's voice. This increase in similarity was prominent in frontal cortical recordings for both the global MMR waveform and the local late (around 850 ms), and early (around 250 ms) latency windows, covering the respective positive and negative potential shifts observed in the ERPs to mother deviant and novel deviant voices at the test session. Intriguingly, the magnitude of the late positive MMR to the novel deviant voice at the test session increased with the number of spindles during the post-familiarization nap. Moreover, higher spindle amplitude was associated with higher MMR similarity between the novel deviant and mother deviant voices in this late latency window, suggesting a role of sleep spindle activity during NonREM sleep in the formation of long-term memory in infants.

4.1. Late positivity in the MMR reflects long-term memory

To assess long-term memory processing, our mismatch voice task used the infant's mother's voice as a second rarely presented stimulus, deviating from the frequently presented standard stimulus. Compared to responses to the standard stimulus, the ERP to the mother deviant stimulus featured a marked positive potential shift over frontal cortical regions starting around 700 ms after stimulus onset. More positive ERPs to rarely presented deviant stimuli compared to frequently presented standard stimuli have been consistently observed in a variety of infant studies (Dehaene-Lambertz & Pena, 2001; Friederici et al., 2007; Friedrich et al., 2009; Friedrich et al., 2004; Jing & Benasich, 2006; Morr et al., 2002; Trainor et al., 2003). Consistent with the view on long-term memory processing being reflected by the positive shift, it was already evident during baseline session for the familiar mother's voice, for which a distinct representation already existed in long-term memory. The positive ERP shift to the mother's voice at baseline could alternatively be explained by an adaptation to the frequent standard stimulus and a novelty effect for the rare deviant stimulus. However, it did not occur in response to the novel deviant voice, which was still unfamiliar at that time and had not yet established its own distinct representation in long-term memory, likely activating the same (voice-independent) representation of the spoken word "baby" as the standard voice. Previous infant research on voice processing has also observed a similar late positive potential shift, showing that long-term memory of a familiar deviant is associated with a more pronounced positive deflection compared to responses to an unfamiliar voice (Beauchemin et al., 2011; Zinke et al., 2018). Unlike these previous studies where the positive shift started around 364 and 300 ms post-stimulus, in our study, the shift emerged somewhat later. This discrepancy might be attributed to minor differences between the studies in the age of the infants, in stimulus duration, discriminability of the voices, or technical processing of the EEG signals (Friederici et al., 2007; Friedrich et al., 2009; Jing & Benasich, 2006; Morr et al., 2002).

Building on this research, we show for the first time that the positive MMR in infants signifies not only established memory representations like the mother voice, but also the formation of a new long-term memory. Specifically, our findings reveal that the MMR to the novel deviant voice lacked the late positive potential shift at the baseline session. However, this shift emerged exclusively at the test session, following the infants' familiarization with this voice through a children's story and subsequent sleep. We demonstrated that at the test session, the MMR to the now-familiar novel deviant stimuli not only included a late positive potential shift but also that this shift closely resembled that observed in response to the mother deviant, both globally across the entire waveform at frontal channels and specifically in the late latency window around 850 ms post-stimulus. This pattern suggests that familiarization followed by sleep facilitated the successful consolidation of the novel deviant voice representation into long-term memory.

It is noteworthy that the late positive MMR predominates over the frontal cortex indicative of a distinctly rapid developmental trajectory in this region (Kolk & Rakic, 2022). Despite its structural immaturity at this age, neuroimaging evidence has identified the prefrontal cortex as a functionally rapidly developing region (for a review, see Hodel, 2018). The pronounced MMR shift over the frontal cortex is also consistent with findings in developing rodents demonstrating that especially medial prefrontal cortex regions like the prelimbic region are causally involved in the formation and retrieval of long-term memory from early life on (e.g., Bock et al., 2014; Contreras et al., 2023; Shan et al., 2022). The prefrontal networks may support the coordinate activation of long-term representations residing in distributed networks in more posterior cortical and subcortical regions.

4.2. Early negativity in the MMR reflects short-term memory processing

In addition to the late positive potential shift, the MMR to both mother deviant voice and novel deviant voice comprised a more transient negative potential shift in an earlier time window around 250 ms post-stimulus. This shift was only observed at test, but not at the baseline session. Similar early negative MMRs have previously been reported in infants, although less consistently than the positive MMRs (Cheour et al., 2002; Friedrich et al., 2009; Jing & Benasich, 2006; Morr et al., 2002; Trainor et al., 2003). This negative infant MMR depends on the strength of the difference between deviant and standard (Morr et al., 2002) and becomes more stable with increasing age (Jing & Benasich, 2006; Trainor et al., 2003). For prosodic characteristics at the word level, the negative MMR did not occur for a non-native language deviant, but only for the native language deviant, where it was related to the individual infant's language outcome two years later (Friedrich et al., 2009).

Interestingly, in a previous study examining infant ERP responses during a single session with the same voice mismatch paradigm as in the present study, we observed an early negative potential shift that appeared to specifically characterize the MMR to the unfamiliar novel deviant voice, suggesting that this early negative MMR is linked to the processing of deviance (from the frequently presented standard stimulus) in short-term memory (Zinke et al., 2018). Diverging from this study, we could not identify such an early negative MMR during the baseline session, although the infants and, accordingly, the data at baseline represented a subsample of that foregoing study. Indeed, the lack of detectability of the early negative MMR in a diminished sample size (from 31 in Zinke et al. to 20 infants in the present study), indicates that the occurrence of such a negative potential shift, at a first presentation of the voice mismatch paradigm, is not a robust phenomenon. However, we did observe such early negative MMR at the second presentation, i.e., the test session of the task, suggesting that the formation of long-term memory for the rarely presented novel voice (as reflected by the late positive MMR) simultaneously supports the occurrence of the early MMR as a reflection of an enhanced processing of the deviance of these stimuli in short-term memory (Beauchemin et al., 2011). This view is supported by our analyses indicating that the larger the late positive MMR to the novel deviant voice was at the test session the larger was the early negative MMR.

Together, these studies support the view of differential memory processes reflected by the early negative and late positive MMR. Although both being maximal over the frontal cortex, their different latency windows, polarity, and occurrence point to qualitatively distinct cognitive processes and underlying neural mechanisms. The late positivity was consistently found to be associated with the activation of long-term memory representations of the voice stimuli, whereas the early negativity appears to be more linked to task-related deviant processing, like the detection of a mismatch between the rarely presented deviant stimuli and the anticipated frequent standard stimulus. This early negative MMR resembles the mismatch negativity observed in adults that has been implicated in the detection of sensory change (Cowan et al., 1993; Titova & Naatanen, 2001).

4.3. Does sleep play a role?

Sleep is recognized as crucial for long-term memory formation in adults, with increasing evidence suggesting even greater importance during early development (Brodt et al., 2023; Gomez & Edgin, 2015; Seehagen, 2019). In both adults and older children, sleep spindles serve as a fundamental mechanism for plastic processes during sleep. Starting from six months of age, spindles have been associated with the formation of new memories (Friedrich et al., 2022; Friedrich et al., 2019, 2020; Friedrich et al., 2015; Friedrich et al., 2017; Kurdziel et al., 2013). Yet, in infants younger than four months, sleep spindles are less mature, though reliably detectable as early as one month of age, with their density and duration increasing from birth to around four months (Kwon et al., 2023). Our study also observed similar age-related increases in spindle parameters, albeit within a narrower age range of 10–18 weeks.

By demonstrating a link between spindle activity and memory formation in three-month-olds, our findings extend previous results from six-month-olds to an even earlier age. A key finding of our study is that the emergence of a late positive MMR to the novel deviant voice, indicating long-term memory formation for this voice, was predicted by the quantity and amplitude of spindles during the nap following voice familiarization. The more spindles an infant displayed during the post-familiarization nap, the stronger was the increase in the late MMR positive potential shift from the baseline to the test session. Furthermore, the higher the spindle amplitude, the greater the increase in similarity between the late positive MMR responses to the novel and mother deviant voices. Notably, we found a consistent link to the frontal cortical late MMR only for spindles identified in frontal cortical recordings, while correlations with spindles identified in central recording channels were not significant. This topographical match of predictive spindle activity and changes in the late MMR across sleep well agrees with findings in older infants (Friedrich, Molle, Born, & Friederici, 2022; Friedrich, Molle, Friederici, & Born, 2019, 2020; Friedrich et al., 2017) and adults (Fernandez & Luthi, 2020; Rasch & Born, 2013) where the formation of specific memories is often linked to more local increases in spindle activity. Finally, we found sleep duration and sleep spindle activity to be selectively associated with the emergence in the late positive MMR to novel voice stimuli, while associations of spindles with the early negative MMR were not significant, corroborating the view that sleep in general (Hanron et al., 2023), and sleep spindles specifically promote plasticity underlying the formation of long-term memory, as reflected here by the emergent late positive MMR.

Our study has clear limitations. We could not establish a wake control condition because all infants fell asleep during the break, and due to ethical concerns, we did not force the infants to stay awake. Consequently, we were unable to determine if sleep is crucial for the observed effects on indicators of long-term memory formation for an unfamiliar voice. Additionally, we could not directly manipulate sleep spindle activity to investigate its causal role in long-term memory formation. Given findings in older children suggesting that long-term memory can also be formed during wakefulness (e.g., Kurz et al., 2023), we cannot exclude that wakefulness likewise supports memory consolidation in infants although in this case through mechanisms different from spindle-related processes that are specific to sleep. In exploratory analyses we added data from two infants who were excluded from the main analyses as they remained awake throughout the post-familiarization sleep period. Including these infants (with a 0-value for respective sleep parameters) in these analyses did not change the observed correlation between frontal spindle count and the late MMR, which is in line with the view that sleep spindle-associated processes are indeed of particular importance for long-term memory formation in infancy.

In addition to these sleep-specific considerations, general methodological limitations may have impacted our findings. Firstly, the relatively modest sample size poses inherent constraints to our correlational approach. Nevertheless, our sample size aligns with comparable studies in early infancy. Another limitation pertains to the absence of noise canceling headphones for the mothers during the task, potentially introducing a bias in maternal responses during the interaction with the infant. Although the mothers were instructed not to talk to their child, we cannot exclude that any facial expressions in response to the stimuli might have influenced our findings. However, this deliberate choice aimed at preserving the natural dynamics of mother–child interactions, enhancing ecological validity. Lastly, the spatial resolution of our sleep and MMR analyses was confined to three frontal and central electrodes, limiting the precision of our findings. Six channels were used mainly for pragmatic reasonsconsidering the challenges posed by the young age of participants. The division into frontal and central channels, though a compromise, reflects a well-established distinction in the sleep and memory literature. Despite these limitations, our results are well in line with a body of evidence in adults, children, and infants up to an age of six months, and thus extend the view of a memory function of sleep spindles to very early infancy as a period characterized by marked dynamics in the development of sleep in general and spindles in particular.

CRediT authorship contribution statement

Lisa Bastian: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Formal analysis, Data curation. **Eva-Maria Kurz:** Writing – review & editing, Writing – original draft, Validation, Software, Formal analysis, Data curation. **Tim Näher:** Writing – review & editing, Software, Methodology, Formal analysis. **Katharina Zinke:** Writing – review & editing, Project administration, Methodology, Investigation, Conceptualization. **Manuela Friedrich:** Writing – review & editing, Validation, Supervision, Data curation. **Jan Born:** Writing – review & editing, Writing – original draft,

Validation, Resources, Project administration, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

I have shared a link to my data in the attach files section.

Acknowledgments

We thank the subjects and their families for participating in the study and Sara Mednick for helpful comments on the manuscript. This study was supported by a grant to J.B. from the Deutsche Forschungsgemeinschaft (Bo 854/18-1) and the European Research Councel (AdG 883098 SleepBalance).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.nlm.2024.107987.

References

Akaike, H. (1973). Maximum likelihood identification of Gaussian autoregressive moving average models. *Biometrika, 60*(2), 255–265. https://doi.org/10.1093/biomet/60.2.255

Alberini, C. M., & Travaglia, A. (2017). Infantile amnesia: A critical period of learning to learn and remember. *The Journal of Neuroscience, 37*(24), 5783–5795. https://doi.org/10.1523/JNEUROSCI.0324-17.2017

Bastian, L., Samanta, A., Ribeiro de Paula, D., Weber, F. D., Schoenfeld, R., Dresler, M., & Genzel, L. (2022). Spindle-slow oscillation coupling correlates with memory performance and connectivity changes in a hippocampal network after sleep. *Human Brain Mapping, 43*(13), 3923–3943. https://doi.org/10.1002/hbm.25893

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*, 1–48. https://doi.org/10.18637/jss.v067.i01

Beauchemin, M., Gonzalez-Frankenberger, B., Tremblay, J., Vannasing, P., Martinez-Montes, E., Belin, P., Beland, R., Francoeur, D., Carceller, A. M., Wallois, F., & Lassonde, M. (2011). Mother and stranger: An electrophysiological study of voice processing in newborns. *Cerebral Cortex, 21*(8), 1705–1711. https://doi.org/10.1093/cercor/bhq242

Berry, R. B., Brooks, R., Gamaldo, C., Harding, S. M., Lloyd, R. M., Quan, S. F., Troester, M. T., & Vaughn, B. V. (2017). AASM Scoring Manual Updates for 2017 (Version 2.4). *J Clin Sleep Med., 13*(5), 665–666.

Bock, J., Poeggel, G., Gruss, M., Wingenfeld, K., & Braun, K. (2014). Infant cognitive training preshapes learning-relevant prefrontal circuits for adult learning: Learning-induced tagging of dendritic spines. *Cerebral Cortex, 24*(11), 2920–2930. https://doi.org/10.1093/cercor/bht148

Brodt, S., Inostroza, M., Niethard, N., & Born, J. (2023). Sleep-A brain-state serving systems memory consolidation. *Neuron, 111*(7), 1050–1075. https://doi.org/10.1016/j.neuron.2023.03.005

Cheour, M., Leppänen, P. H., & Kraus, N. (2000). Mismatch negativity (MMN) as a tool for investigating auditory discrimination and sensory memory in infants and children. *Clinical Neurophysiology, 111*(1), 4–16. https://doi.org/10.1016/s1388-2457(99)00191-1

Cheour, M., Martynova, O., Naatanen, R., Erkkola, R., Sillanpaa, M., Kero, P., Raz, A., Kaipio, M. L., Hiltunen, J., Aaltonen, O., Savela, J., & Hamalainen, H. (2002). Speech sounds learned by sleeping newborns. *Nature, 415*(6872), 599–600. https://doi.org/10.1038/415599b

Clawson, B. C., Durkin, J., & Aton, S. J. (2016). Form and function of sleep spindles across the lifespan. *Neural plasticity, 2016*(1), 6936381. https://doi.org/10.1155/2016/6936381

Contreras, M. P., Fechner, J., Born, J., & Inostroza, M. (2023). Accelerating maturation of spatial memory systems by experience: evidence from sleep oscillation signatures of memory processing. *The Journal of Neuroscience, 43*(19), 3509–3519. https://doi.org/10.1523/JNEUROSCI.1967-22.2023

Cowan, N., Winkler, I., Teder, W., & Naatanen, R. (1993). Memory prerequisites of mismatch negativity in the auditory event-related potential (ERP). *Journal of Experimental Psychology. Learning, Memory, and Cognition, 19*(4), 909–921. https://doi.org/10.1037//0278-7393.19.4.909

DeCasper, A. J., & Fifer, W. P. (1980). Of human bonding: Newborns prefer their mothers' voices. *Science, 208*(4448), 1174–1176. https://doi.org/10.1126/science.7375928

Dehaene-Lambertz, G., & Pena, M. (2001). Electrophysiological evidence for automatic phonetic processing in neonates. *Neuroreport, 12*(14), 3155–3158. https://doi.org/10.1097/00001756-200110080-00034

Diekelmann, S., & Born, J. (2010). The memory function of sleep. *Nature Reviews Neuroscience, 11*(2), 114–126. https://doi.org/10.1038/nrn2762

Donoghue, T., Haller, M., Peterson, E. J., Varma, P., Sebastian, P., Gao, R., Noto, T., Lara, A. H., Wallis, J. D., Knight, R. T., Shestyuk, A., & Voytek, B. (2020). Parameterizing neural power spectra into periodic and aperiodic components. *Nature Neuroscience, 23*(12), 1655–1665. https://doi.org/10.1038/s41593-020-00744-x

Fernandez, L. M. J., & Luthi, A. (2020). Sleep spindles: Mechanisms and functions. *Physiological Reviews, 100*(2), 805–868. https://doi.org/10.1152/physrev.00042.2018

Friederici, A. D., Friedrich, M., & Christophe, A. (2007). Brain responses in 4-month-old infants are already language specific. *Current Biology, 17*(14), 1208–1211. https://doi.org/10.1016/j.cub.2007.06.011

Friedrich, M., Herold, B., & Friederici, A. D. (2009). ERP correlates of processing native and non-native language word stress in infants with different language outcomes. *Cortex, 45*(5), 662–676. https://doi.org/10.1016/j.cortex.2008.06.014

Friedrich, M., Molle, M., Born, J., & Friederici, A. D. (2022). Memory for nonadjacent dependencies in the first year of life and its relation to sleep. *Nature Communications, 13*(1), 7896. https://doi.org/10.1038/s41467-022-35558-x

Friedrich, M., Molle, M., Friederici, A. D., & Born, J. (2019). The reciprocal relation between sleep and memory in infancy: Memory-dependent adjustment of sleep spindles and spindle-dependent improvement of memories. *Developmental Science, 22*(2), e12743. https://doi.org/10.1111/desc.12743

Friedrich, M., Molle, M., Friederici, A. D., & Born, J. (2020). Sleep-dependent memory consolidation in infants protects new episodic memories from existing semantic memories. *Nature Communications, 11*(1), 1298. https://doi.org/10.1038/s41467-020-14850-8

Friedrich, M., Weber, C., & Friederici, A. D. (2004). Electrophysiological evidence for delayed mismatch response in infants at-risk for specific language impairment. *Psychophysiology, 41*(5), 772–782. https://doi.org/10.1111/j.1469-8986.2004.00202.x

Friedrich, M., Wilhelm, I., Born, J., & Friederici, A. D. (2015). Generalization of word meanings during infant sleep. *Nature Communications, 6*, 6004. https://doi.org/10.1038/ncomms7004

Friedrich, M., Wilhelm, I., Mölle, M., Born, J., & Friederici, A. D. (2017). The sleeping infant brain anticipates development. *Current Biology, 27*(15), 2374–2380. https://doi.org/10.1016/j.cub.2017.06.070

Gais, S., Molle, M., Helms, K., & Born, J. (2002). Learning-dependent increases in sleep spindle density. *The Journal of Neuroscience, 22*(15), 6830–6834. https://doi.org/10.1523/JNEUROSCI.22-15-06830.2002

Gomez, R. L., & Edgin, J. O. (2015). Sleep as a window into early neural development: Shifts in sleep-dependent learning effects across early childhood. *Child Development Perspectives, 9*(3), 183–189. https://doi.org/10.1111/cdep.12130

Hanron, O., Mason, G. M., Holmes, J. F., & Spencer, R. M. C. (2023). Early childhood naps initiate emotional memory processing in preparation for enhanced overnight consolidation. *Child Development, 94*(3), 721–733. https://doi.org/10.1111/cdev.13890

Hayne, H., Gross, J., Hildreth, K., & Rovee-Collier, C. (2000). Repeated reminders inrease the speed of memory retrieval by 3-month-old infants. *Developmental Science, 3*(3), 312–318. https://doi.org/10.1111/1467-7687.00125

Hodel, A. S. (2018). Rapid infant prefrontal cortex development and sensitivity to early environmental experience. *Developmental Review, 48*, 113–144. https://doi.org/10.1016/j.dr.2018.02.003

Hoehl, S., & Wahl, S. (2012). Recording infant ERP data for cognitive research. *Developmental Neuropsychology, 37*(3), 187–209. https://doi.org/10.1080/87565641.2011.627958

Horvath, K., Hannon, B., Ujma, P. P., Gombos, F., & Plunkett, K. (2018). Memory in 3-month-old infants benefits from a short nap. *Developmental Science, 21*(3), e12587. https://doi.org/10.1111/desc.12587

Huber, R., & Born, J. (2014). Sleep, synaptic connectivity, and hippocampal memory during early development. *Trends in Cognitive Sciences, 18*(3), 141–152. https://doi.org/10.1016/j.tics.2013.12.005

Jing, H., & Benasich, A. A. (2006). Brain responses to tonal changes in the first two years of life. *Brain & Development, 28*(4), 247–256. https://doi.org/10.1016/j.braindev.2005.09.002

Klinzing, J. G., Niethard, N., & Born, J. (2019). Mechanisms of systems memory consolidation during sleep. *Nature Neuroscience, 22*(10), 1598–1610. https://doi.org/10.1038/s41593-019-0467-3

Kolk, S. M., & Rakic, P. (2022). Development of prefrontal cortex. *Neuropsychopharmacology, 47*(1), 41–57. https://doi.org/10.1038/s41386-021-01137-9

Kurdziel, L., Duclos, K., & Spencer, R. M. (2013). Sleep spindles in midday naps enhance learning in preschool children. *Proceedings of the National Academy of Sciences of the United States of America, 110*(43), 17267–17272. https://doi.org/10.1073/pnas.1306418110

Kurz, E.-M., Zinke, K., & Born, J. (2023). Sleep electroencephalogram (EEG) oscillations and associated memory processing during childhood and early adolescence. *Developmental Psychology, 59*(2), 297–311. https://doi.org/10.1037/dev0001487

Kushnerenko, E., Winkler, I., Horvath, J., Naatanen, R., Pavlov, I., Fellman, V., & Huotilainen, M. (2007). Processing acoustic change and novelty in newborn infants.

The European Journal of Neuroscience, 26(1), 265–274. https://doi.org/10.1111/j.1460-9568.2007.05628.x

Kushnerenko, E. V., Van den Bergh, B. R., & Winkler, I. (2013). Separating acoustic deviance from novelty during the first year of life: A review of event-related potential evidence. *Frontiers in Psychology, 4*, 595. https://doi.org/10.3389/fpsyg.2013.00595

Kwon, H., Walsh, K. G., Berja, E. D., Manoach, D. S., Eden, U. T., Kramer, M. A., & Chu, C. J. (2023). Sleep spindles in the healthy brain from birth through 18 years. *Sleep, 46*(4), zsad017. https://doi.org/10.1093/sleep/zsad017

Latchoumane, C. F. V., Ngo, H. V. V., Born, J., & Shin, H. S. (2017). Thalamic spindles promote memory formation during sleep through triple phase-locking of cortical, thalamic, and hippocampal rhythms. *Neuron, 95*(2), 424–435. https://doi.org/10.1016/j.neuron.2017.06.025

Leppanen, P. H., Guttorm, T. K., Pihko, E., Takkinen, S., Eklund, K. M., & Lyytinen, H. (2004). Maturational effects on newborn ERPs measured in the mismatch negativity paradigm. *Experimental Neurology, 190*(Suppl 1), S91–S101. https://doi.org/10.1016/j.expneurol.2004.06.002

Lüdecke, D., Ben-Shachar, M. S., Patil, I., & Makowski, D. (2020). Extracting, computing and exploring the parameters of statistical models using R. *The Journal of Open Source Software, 5*(53), 2445. https://doi.org/10.21105/joss.02445

Mai, X., Xu, L., Li, M., Shao, J., Zhao, Z., deRegnier, R. A., Nelson, C. A., & Lozoff, B. (2012). Auditory recognition memory in 2-month-old infants as assessed by event-related potentials. *Developmental Neuropsychology, 37*(5), 400–414. https://doi.org/10.1080/87565641.2011.650807

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

Mason, G. M., Lokhandwala, S., Riggins, T., & Spencer, R. M. C. (2021). Sleep and human cognitive development. *Sleep Medicine Reviews, 57*, Article 101472. https://doi.org/10.1016/j.smrv.2021.101472

Morr, M. L., Shafer, V. L., Kreuzer, J. A., & Kurtzberg, D. (2002). Maturation of mismatch negativity in typically developing infants and preschool children. *Ear and Hearing, 23*(2), 118–136. https://doi.org/10.1097/00003446-200204000-00005

Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J. M. (2011). FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Computational Intelligence and Neuroscience, 2011*, Article 156869. https://doi.org/10.1155/2011/156869

Partanen, E., Kujala, T., Naatanen, R., Liitola, A., Sambeth, A., & Huotilainen, M. (2013). Learning-induced neural plasticity of speech processing before birth. *Proceedings of the National Academy of Sciences of the United States of America, 110*(37), 15145–15150. https://doi.org/10.1073/pnas.1302159110

Rasch, B., & Born, J. (2013). About sleep's role in memory. *Physiological Reviews, 93*(2), 681–766. https://doi.org/10.1152/physrev.00032.2012

Schwarz, G. (1978). Bayesian information criterion. *The Annals of Statistics, 6*, 461–464.

Seehagen, S. Z., & Herbert, J. S. (2019). Remembering in the context of internal states: The role of sleep for infant memory. *Child Development Perspectives, 13*(2), 110–115. https://doi.org/10.1111/cdep.12321

Shan, X., Contreras, M. P., Mendez, M., Born, J., & Inostroza, M. (2022). Unfolding of spatial representation at systems level in infant rats. *Hippocampus, 32*(2), 121–133. https://doi.org/10.1002/hipo.23392

Timofeev, I., Grenier, F., Bazhenov, M., Houweling, A. R., Sejnowski, T. J., & Steriade, M. (2002). Short- and medium-term plasticity associated with augmenting responses in cortical slabs and spindles in intact cortex of cats in vivo. *The Journal of Physiology, 542*(Pt 2), 583–598. https://doi.org/10.1113/jphysiol.2001.013479

Titova, N., & Naatanen, R. (2001). Preattentive voice discrimination by the human brain as indexed by the mismatch negativity. *Neuroscience Letters, 308*(1), 63–65. https://doi.org/10.1016/s0304-3940(01)01970-x

Trainor, L., McFadden, M., Hodgson, L., Darragh, L., Barlow, J., Matsos, L., & Sonnadara, R. (2003). Changes in auditory cortex and the development of mismatch negativity between 2 and 6 months of age. *International Journal of Psychophysiology, 51*(1), 5–15. https://doi.org/10.1016/s0167-8760(03)00148-x

Zinke, K., Thone, L., Bolinger, E. M., & Born, J. (2018). Dissociating long and short-term memory in three-month-old infants using the mismatch response to voice stimuli. *Frontiers in Psychology, 9*, 31. https://doi.org/10.3389/fpsyg.2018.00031