

Beat processing in newborn infants cannot be explained by statistical learning based on transition probabilities

Gábor P. Hádén^{a,b,*}, Fleur L. Bouwer^{c,d,e,f}, Henkjan Honing^{c,d}, István Winkler^a

^a Institute of Cognitive Neuroscience and Psychology, HUN-REN Research Centre for Natural Sciences, Magyar tudósok körútja 2, H-1117 Budapest, Hungary

^b Department of Telecommunications and Media Informatics, Faculty of Electrical Engineering and Informatics, Budapest University of Technology and Economics, Magyar tudósok körútja 2, 1117 Budapest, Hungary

^c Music Cognition Group, Institute for Logic, Language, and Computation, University of Amsterdam, P.O. Box 94242, 1090 GE Amsterdam, the Netherlands

^d Amsterdam Brain and Cognition, University of Amsterdam, P.O. Box 15900, 1001 NK Amsterdam, the Netherlands

^e Department of Psychology, Brain & Cognition, University of Amsterdam, P.O. Box 15900, 1001 NK Amsterdam, the Netherlands

^f Cognitive Psychology Unit, Institute of Psychology & Leiden Institute for Brain and Cognition, Leiden University, 2333 AK Leiden, the Netherlands

ARTICLE INFO

Keywords:

Neonate
EEG
Rhythm
Beat perception
Statistical learning

ABSTRACT

Newborn infants have been shown to extract temporal regularities from sound sequences, both in the form of learning regular sequential properties, and extracting periodicity in the input, commonly referred to as a regular pulse or the 'beat'. However, these two types of regularities are often indistinguishable in isochronous sequences, as both statistical learning and beat perception can be elicited by the regular alternation of accented and unaccented sounds. Here, we manipulated the isochrony of sound sequences in order to disentangle statistical learning from beat perception in sleeping newborn infants in an EEG experiment, as previously done in adults and macaque monkeys. We used a binary accented sequence that induces a beat when presented with isochronous timing, but not when presented with randomly jittered timing. We compared mismatch responses to infrequent deviants falling on either accented or unaccented (i.e., odd and even) positions. Results showed a clear difference between metrical positions in the isochronous sequence, but not in the equivalent jittered sequence. This suggests that beat processing is present in newborns. Despite previous evidence for statistical learning in newborns the effects of this ability were not detected in the jittered condition. These results show that statistical learning by itself does not fully explain beat processing in newborn infants.

1. Introduction

Processing sound sequences is crucial for both music and speech perception (Patel, 2008; Kotz, Ravignani, & Fitch, 2018). Studying the development of sound processing in infants can provide insights into the bootstrapping of speech and music acquisition during a period in which learning and brain maturation go hand in hand (Finlay & Uchiyama, 2020). In sound sequences, several types of structure can be discerned that contribute to efficient processing, including regularities in the temporal aspects of a sequence (i.e., the 'beat'), regularities in order of events (i.e., statistical regularities), and hierarchical temporal stress patterns (i.e., the meter; Langus, Mehler, & Nespore, 2017). While both the processing of temporal (Winkler, Hádén, Ladinig, Sziller, & Honing, 2009) and statistical (Bosseler, Teinonen, Tervaniemi, & Huotilainen,

2016) regularities have been suggested to be present in neonates, the relationship between these remains unclear, since sequences used to probe statistical regularities often also contain temporal regularities and vice versa (Bouwer, Van Zuijlen, & Honing, 2014). Here we aim to disentangle the ability of neonates to detect regularities in time (the beat or pulse), and regularities in order (statistical learning).

Temporal structure in rhythm often comes in the form of a regular beat: a series of perceived regularly recurring salient events (Cooper & Meyer, 1963; Honing, 2012; Honing & Bouwer, 2019). The ability to synchronize and coordinate movement with a regular beat is called rhythmic entrainment. This coordination is thought to result from the coupling of internal low frequency oscillations in auditory and motor areas with the external rhythmic signal (neural entrainment; see Large, Herrera, & Velasco, 2015). Beat-based timing has been suggested to be

* Corresponding author at: Institute of Cognitive Neuroscience and Psychology, HUN-REN Research Centre for Natural Sciences, Magyar tudósok körútja 2, Budapest H-1117, Hungary.

E-mail addresses: haden.gabor@ttk.hu (G.P. Hádén), f.l.bouwer@fsw.leidenuniv.nl (F.L. Bouwer), honing@uva.nl (H. Honing), winkler.istvan@ttk.hu (I. Winkler).

<https://doi.org/10.1016/j.cognition.2023.105670>

Received 27 March 2023; Received in revised form 8 November 2023; Accepted 17 November 2023

Available online 27 November 2023

0010-0277/© 2023 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC license (<http://creativecommons.org/licenses/by-nc/4.0/>).

somewhat separate from other timing processes, like the perception of (a sequence of) single absolute temporal intervals (Bouwer, Fahrenfort, Millard, Kloosterman, & Slagter, 2023; Bouwer, Honing, & Slagter, 2020; Breska & Deouell, 2017; Honing & Merchant, 2014; Teki, Grube, Kumar, & Griffiths, 2011). Note that we will refer to the detection of beat as shown by brain responses as “beat perception” to keep the terminology compatible with previous literature, although sleeping infants cannot describe their experience and no overt behavior is involved here, making this term incorrect in the literal. In humans, we have previously suggested that beat perception is already functional at birth (Winkler et al., 2009). However, the results of this study could have been biased by comparing between responses to acoustically different sounds and contexts (detailed in Bouwer et al., 2014). Hence, the presence of beat perception at birth remains to be confirmed.

Specifically, beat perception needs to be dissociated from the detection of statistical regularities in the order of events, such as transitional/conditional probabilities of item succession, as was described for language-like (Saffran, Newport, & Aslin, 1996) and in tonal stimuli (Saffran, Johnson, Aslin, & Newport, 1999). Statistical learning in its

narrowest sense is the extraction of these regularities from the order of the elements of the input without explicit feedback or even awareness after prolonged exposure (Conway, 2020). Statistical learning is thought to be a domain general mechanism manifest in different domains, including audition (Saffran et al., 1996), music (Pearce, 2018; Saffran et al., 1999), and vision (Duncan & Theeuwes, 2020; Fiser & Aslin, 2001), and most prominently in language. There is evidence that statistical learning is already functional at birth (Bosseler et al., 2016; Bulf, Johnson, & Valenza, 2011; Teinonen, Fellman, Näättänen, Alku, & Huotilainen, 2009). However, there is an ongoing debate whether statistical learning in children is an unchanging domain general mechanism or it shows developmental changes that differ between the visual and the auditory modality (e.g., Saffran, Newport, Aslin, Tunick, & Barrueco, 1997; Raviv & Arnon, 2018; reviewed in Conway, 2020).

Understanding the processing of sound sequences in infants requires one to dissociate the processing of the beat, absolute temporal intervals, and statistical regularities in the order of sounds. Firstly, to show beat perception, the rhythmic stimulus needs to be acoustically and/or temporally varied, so that these can be distinguished from interval-

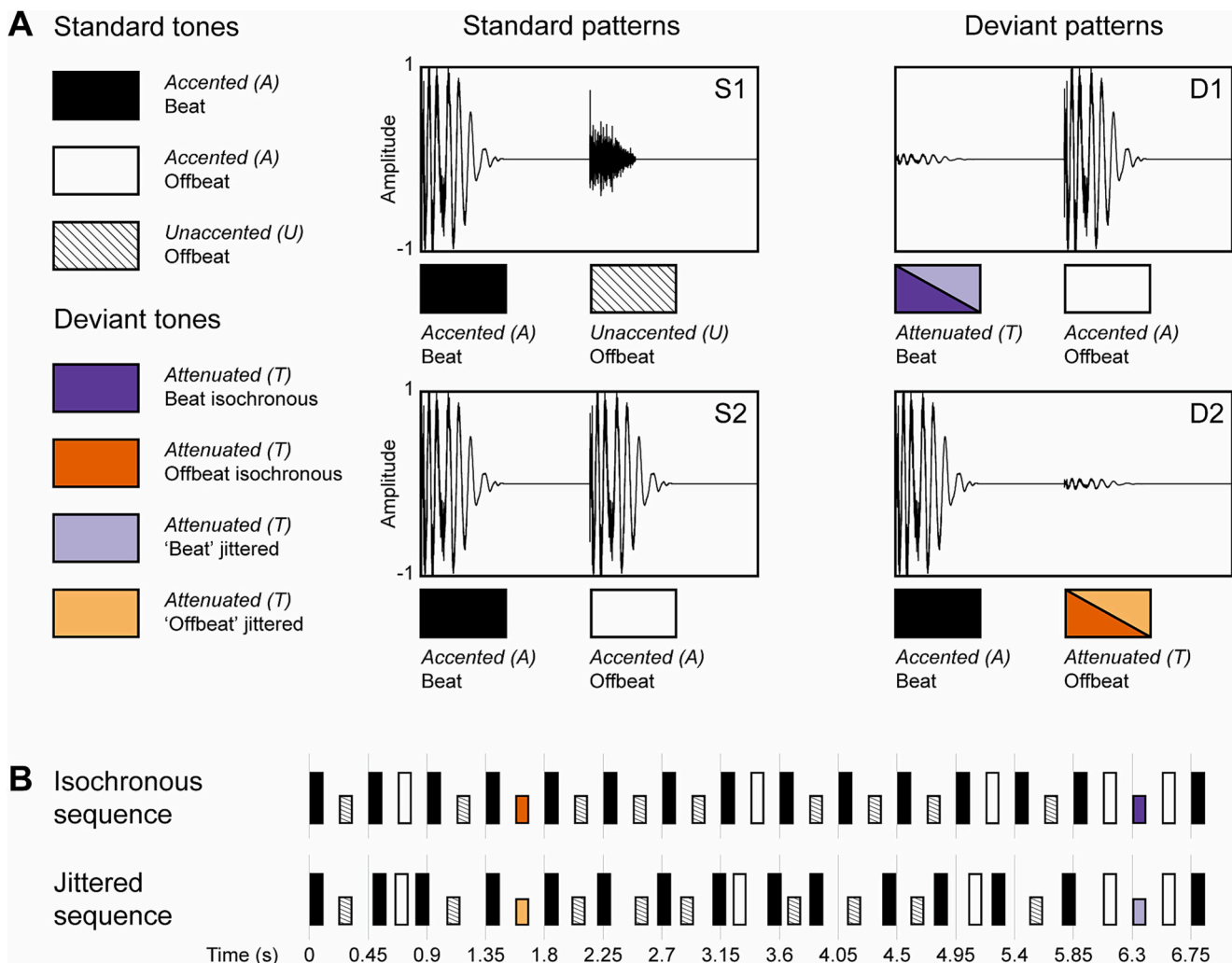


Fig. 1. Schematic diagram of the rhythmic stimulus patterns used in the experiment. (A) The two standard (S1 and S2) and two deviant patterns (D1 and D2) are made up of three different sounds (A = accented, U = unaccented, and T = attenuated). An accented sound could occur either on the beat or offbeat, an unaccented sound was restricted to the offbeat position. Attenuated sounds were used as deviants in both positions (beat and offbeat) and conditions (isochronous and jittered). (B) Standard and deviant sound patterns were concatenated into a single rhythmic stream in a random order (see main text for details). Sequences in the isochronous condition had an inter-onset interval (IOI) of 225 ms, in the jittered condition these were randomly chosen from the range 150 to 300 ms using a uniform distribution. Deviants were always preceded and followed by an accented sound, with a fixed IOI of 225 ms in both conditions. Note that while we use the labels Beat and Offbeat in both conditions to refer to the position of the deviants in the rhythmic sequence, but a Beat (or Offbeat) can, of course, only be sensed in the isochronous condition, and not in the jittered condition. (Adapted from Honing et al., 2018).

based perception (Bouwer et al., 2020; Bouwer, Werner, Knetemann, & Honing, 2016; Honing, Bouwer, Prado, & Merchant, 2018). Secondly, while the presence of acoustic variation in a sequence (e.g., differences in pitch or intensity between events) can aid the detection of a beat, especially in musical novices (Bouwer, Burgoyne, Odijk, Honing, & Grahn, 2018), it also creates possible confounds when probing beat perception, as differences in responses related to the temporal structure (e.g., different responses to events on and off the beat) may in fact be caused by differences in acoustic and sequential order properties when using acoustically rich and varied rhythmic sequences (Bouwer et al., 2014). Thus, to show beat perception, it must be carefully dissociated from learning transitional probabilities.

To disentangle beat perception from statistical learning and the perception of interval-based temporal structure (e.g., learning an absolute interval), an auditory oddball paradigm was proposed by Bouwer et al. (2016), also used in Honing et al. (2018). The paradigm employs a rhythmic sequence made up of a pattern of loud (on all odd, “beat” positions and, a small portion of even, “offbeat” positions) and soft (in most even, “offbeat” positions) percussive sounds (Fig. 1A), such that the acoustic stimulus could induce a simple binary metrical structure (“duple meter”). The presence of timbre and intensity differences arguably creates an ecological way of inducing a beat (Ladinig, Honing, Håden, & Winkler, 2009). Within the context of this clearly beat-inducing sequence of alternating loud and soft sounds, in a proportion of the patterns, the offbeat positions are also filled with loud sounds. These patterns are used to probe beats and offbeats with identical acoustic properties. These rhythmic sequences are presented in two conditions: an isochronous condition, and a jittered condition (Fig. 1B). In the isochronous condition, sounds are presented with a constant inter-onset interval (IOI), allowing a beat to be induced (i.e., one metrical level of a duple meter). In the jittered condition, the IOIs are irregular (i.e., not isochronous), thus disabling the perception of a regular beat. However, the jittered sequences still contain the same order-based statistical regularity (alternation) between the louder and softer sounds as the isochronous ones. As such, the jittered condition serves as a control for statistical learning of the succession of sounds with different timbres.

Bouwer et al. (2016) presented these sequences to adults, and measured event-related potentials (ERPs) to rare, unexpected intensity decrements (i.e., deviants) to assess the formation of beat-based expectations. When the prediction of incoming stimuli based on previous stimuli fails, that is, the regularity extracted from the sound sequence is violated, the mismatch negativity (MMN) ERP component is elicited in adults (for a recent review, see Fitzgerald & Todd, 2020). The amplitude of the MMN increases together with the specificity of the prediction (Southwell & Chait, 2018) and the amount of deviation from the predicted sound (e.g., Novitski, Tervaniemi, Huotilainen, & Näätänen, 2004).

Bouwer et al. (2016) found larger MMN responses for deviants in odd (beat) than even (offbeat) positions, suggesting that the prediction was stronger on the odd position. This difference between predictions for beat and offbeat positions was only evident in the isochronous condition under unattended conditions, in which the MMN was measured. The authors argue that this was due to beat perception contributing to differences between positions only in the isochronous, but not the jittered condition because in the latter, no beat should emerge. Of note, in the jittered condition, significant difference in the amplitude between beat and offbeat positions did appear in the P3a range under unattended conditions, as well as in the N2 range when the sounds were attended. These ERP amplitude differences suggest that adults picked up on the statistical differences between beat and offbeat positions even in jittered sequences under both attended and unattended conditions, albeit to a lesser extent than when the sequences were isochronous and beat perception additionally contributed to position differences.

Moreover, in both beat and offbeat positions, the P3a (an ERP component indexing post-processing following the detection of an unattended regularity violation; for a review, see Polich, 2007) was larger

in the isochronous than the jittered condition, indicating that the isochrony of the sequence increased the amount of processing needed after a regularity violation. This held true for both attended and unattended blocks.

A mismatch response (MMR) similar to MMN can be measured in young infants in response to regularity violations, including sleeping neonates (Alho, Sainio, Sajaniemi, Reinikainen, & Näätänen, 1990; Håden, Németh, Török, & Winkler, 2016). In our previous study in newborn infants (Winkler et al., 2009) stimulus omissions (Håden, Honing, Török, & Winkler, 2015) at different metrical positions of a repetitive rhythmic sequence were used to test beat and meter processing. However, as was noted before, the different metrical positions had different acoustic and sequential properties. This makes it possible that differences in omission MMR responses found for different metrical positions were based on learning the statistical properties of the order of different tones, that is, by statistical learning (Bouwer et al., 2014).

Here, we presented a variant of Bouwer et al. (2016) stimulus paradigm (same stimuli, but no active condition, no silent movie presented, fewer but longer stimulus blocks, cf. Procedure sections) to sleeping newborn infants to separate the effects of beat perception from those of statistical learning (Fig. 1) in order to provide converging evidence to the notion that newborns brains can process the beat in rhythm, as suggested in our previous study (Winkler et al., 2009), while controlling for the effects of statistical learning. Response differences (MMR) between rare deviant and corresponding standard sounds were calculated at beat (odd) and offbeat (even) positions, separately for the isochronous and the jittered sequences (note that we refer to odd and even positions as “beat” and “offbeat” even for the jittered sequences, to make the terminology more consistent). In line with the effects in adults, firstly we expected newborns to learn the statistically predictable alternation between louder and softer sounds. Since beat perception does not occur in the jittered condition and the MMR is calculated by comparing the responses to sounds with identical acoustic properties and context (i.e., no acoustic difference between the tones in the compared nor in the preceding position), differences in the MMR responses between beat and offbeat positions should result from learning the sequential statistical regularities of the sound sequence (i.e., the alternation). Secondly, we expected to find evidence for beat perception, to corroborate our previous results (c.f. Bouwer et al., 2016 as well as Winkler et al., 2009). Based on Bouwer et al. (2016) results, beat perception should make the MMR difference between beat and offbeat position larger in the isochronous than in the jittered condition, as it would additionally contribute to the beat-offbeat difference.

2. Material and methods

2.1. Participants

EEG was recorded from 31 (12 male) healthy, full-term newborn infants during day 0–6 (mean = 2.06, SD = 1.01) postpartum. The mean gestational age was 39.31 weeks (SD = 0.87), mean birth weight was 3319 g (SD = 311.80). All participants had an Apgar score of 9/10 (1 min./5 min.). Due to excessive electrical artifacts, 4 participants with <50% of the trials retained for any of the tested stimuli were excluded from the analysis. The final sample consisted of 27 (12 male) participants. Infants were predominantly asleep during the recordings, in quiet sleep 87% of the time and in active sleep 11% of the time (Anders, Emde, & Parmelee, 1971). Informed consent was obtained from the mother or both parents. The experiment was carried out in a dedicated experimental room.

2.2. Stimuli

The stimuli were identical to those used in Bouwer et al. (2016) and Honing et al. (2018), and the following description was adapted from Honing et al. (2018). See Fig. 1 for a schematic depiction of the rhythmic

stimulus patterns used, Fig. 2 for a depiction of the stimuli generation as a transition network, and Supplementary Materials for audio examples of the two conditions.

To induce a simple binary metrical structure (duple meter) two sounds differing in timbre, intensity and duration were used. The sounds were made with QuickTime's (Apple, Inc., Cupertino, CA, U.S.A.) drum timbres. The first sound (hereafter Accented or A) consisted of simultaneous sounding of bass drum and hi-hat, whereas the second sound (hereafter Unaccented or U) consisted of a hi-hat sound. Sound lengths were 110 ms and 70 ms, respectively. A third, deviant sound (hereafter Attenuated or T), was created by attenuating the accented sound by 25 dB using Praat software (www.praat.org).

A rhythmic sound sequence was created by concatenating patterns consisting of two sounds. Sixty percent of the patterns consisted of an accented sound followed by an unaccented sound (S1, used to create a context with a clear duple meter) and 30% of the patterns consisted of an accented sound followed by another accented sound (S2, used to create offbeats with an accented sound, to allow deviants on the beat to have the same acoustic context as deviants off the beat). In the remaining 10% of the sequence either an attenuated sound followed an accented sound (5%; D1, deviant in even position, offbeat) or an accented sound followed an attenuated sound (5%; D2, deviant in odd position, beat). Concatenating the four patterns in a random order resulted in accented and unaccented sounds alternating most of the time.

The resulting sequence was used in both the isochronous and jittered conditions. In the isochronous condition, all sounds were presented with a constant inter-onset interval of 225 ms. In this condition, the alternation of accented and unaccented sounds was expected to induce a beat (or duple meter) with an inter-beat interval of 450 ms, which is within the optimal range for beat perception in humans (London, 2002). Sounds in odd positions of the sequence (including deviant D1) can be considered on the beat, while sounds in even positions (including deviant D2) offbeat. In the jittered condition, the inter-onset intervals were randomly distributed between 150 and 300 ms with an average of 225 ms (uniform distribution). This resulted in a sequence with the same order of stimuli as the isochronous one, while not inducing a regular beat (Honig, 2012; London, 2002) (note that D1 and D2 will be referred to as on the beat or offbeat for both conditions, although perceptually this is only true for the isochronous condition).

Sequences were created with the following five constraints. The inter-onset interval just before and after a deviant tone was kept constant at 225 ms. A deviant on the beat (D1) was always preceded by an

accented sound offbeat (S2). These two constraints made the acoustic and preceding temporal context of the deviants identical between positions and conditions. To optimize the possibility of inducing beat in the isochronous condition, S2 (containing two consecutive accented sounds) was never repeated, and only a maximum of four consecutive S1 patterns was allowed. Finally, at least five standard patterns (S1 and S2) were delivered between two consecutive deviant patterns.

An alternative description of the rhythmic sound sequence generation (at the level of the individual sounds A, U and T) is given in Fig. 2. This transition network shows the transition probabilities between the three individual sounds used in both the isochronous and jittered conditions.

2.3. Procedure

Stimuli were presented in 6 blocks with 3 blocks consisting of isochronous and 3 blocks of jittered sequences (1300 patterns, i.e., 2600 sound events, per block). One block lasted 9 min and 45 s, and blocks were separated by a silent interval of at least 15 s. Isochronous and jittered blocks were presented in semi-random order, with a maximum of two blocks from the same condition following each other. This resulted in a total session length of about 60 min.

Sounds were presented binaurally using the E-Prime stimulus presentation software (version 1.1.4.1, Psychology Software Tools, Inc., Pittsburgh, PA, USA) via ER-1 headphones (Etymotic Research Inc., Elk Grove Village, IL, USA) connected via sound tubes to self-adhesive ear cups (Sanibel Supply, Middelfart, Denmark) placed over the infants' ears. The sound intensity was set to a comfortable level for an adult listener. The output of the sound system was hard limited to 80 dB SPL using a purpose built and calibrated circuit between the headphones and computer.

Overall, the design of the study was identical to the unattended condition presented in Bouwer et al. (2016), except that the adult participants watched a silent video with subtitles, whereas infants were mostly asleep (see above.)

2.4. EEG recording and analysis

EEG was recorded with Ag/AgCl electrodes attached to the Fp1, Fp2, Fz, F3, F4, F7, F8, T3, T4, Cz, C3, C4, Pz, P3, P4 locations according to the International 10–20 system, and near to the outer canthus of the right eye. The reference electrode was placed on the tip of the nose and

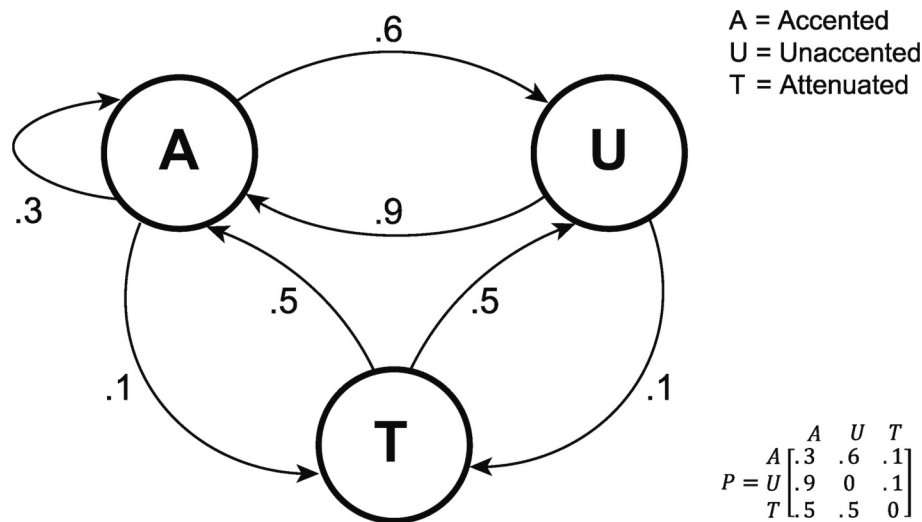


Fig. 2. Rhythmic sequence generation visualized as a transition network, showing the transition probabilities between the individual sound elements A, U and T. Note that this representation is a simplification as it does not include certain constraints on the way concatenated patterns can follow each other (see main text). (Adapted from Honig et al., 2018).

the ground electrode on the forehead. Data was recorded using a direct-coupled amplifier (V-Amp, Brain Products, Munich, Germany) at 24-bit resolution with a sampling rate of 1000 Hz. Signals were off-line filtered between 1.5 and 30 Hz and epochs of –100–500 ms with respect to sound onset were extracted for each of the D1, D2, S1 (only after S2 to control for acoustic context; for brevity referred to simply as S1 from this point on), and S2 stimuli. The 100-ms pre-stimulus interval served as the baseline for amplitude measurements and illustrations. Due to low signal-to-noise ratio on other electrodes only data from electrodes F3,

Fz, F4, C3, Cz, C4, P3, Pz, and P4 were further analyzed. Matlab (version 8.3, Mathworks, Inc., Nantick, MA, USA), EEGLAB (version 2019, Delorme & Makeig, 2004), and ERPLAB (version 7.0, Lopez-Calderon & Luck, 2014) were used for data processing. Epochs with a voltage change exceeding 100 μ V on any channels measured in a moving window (window length = 100 ms, window step = 50 ms) for the entire epoch duration were rejected from the analysis. The mean percentage of accepted trials was 83 (SD = 12.66, for details per condition See Supplementary Table 1) over all conditions. Due to

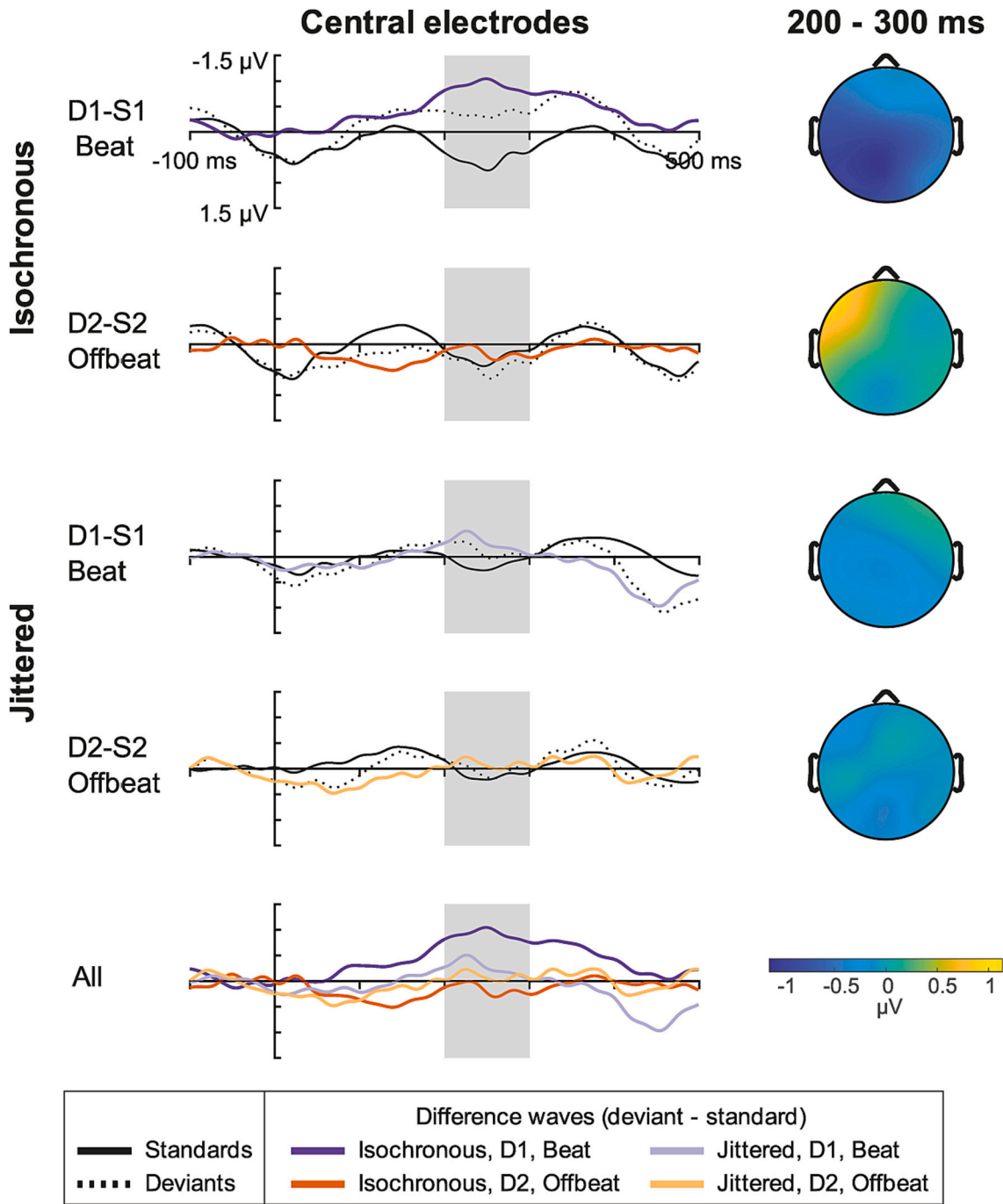


Fig. 3. Left: Grand average (N = 27) ERP responses averaged over the three central electrodes (C3, Cz, C4) to deviant (D1, D2), standard (S1, S2) stimuli and the deviant minus standard difference waveforms (D1-S1, D2-S2) for the isochronous and the jittered condition. The panel on the bottom overplots all four difference waveforms. Right: Scalp distributions the average deviant-minus-standard difference amplitudes from the 200–300 ms time window (marked by grey rectangles on the signals on the left). 0 ms denotes the start of the sound at the Beat position for D1 and S1 and the start of the sound at the Offbeat position for D2 and S2. Note that the labels Beat and Offbeat are used to refer to the position of the deviants in the rhythmic sequence. But a Beat (or Offbeat) can, of course, only be sensed in the isochronous condition, and not in the jittered condition.

experimenter error two participants were presented with only 5 blocks (1 isochronous and 1 jittered block left out, respectively). They were not rejected from the analysis, as the percentage of accepted epochs was over 50 in both conditions, and their lowest number of accepted epochs was 99 and 69 for any condition, respectively.

D1-S1 and D2-S2 difference waveforms were created for both conditions. Mean difference amplitudes were measured from the 200–300 ms time window, relative to stimulus onset (MMR). The window size and latency were determined by visual inspection.

Effects of the two conditions and two positions on MMR amplitudes were tested with a general linear model (GLM) of the structure Timing [Isochronous vs. Jittered] \times Position [Beat vs. Offbeat] \times Frontality [F vs. C vs. P electrode line] \times Laterality [left vs. midline vs. right]. Greenhouse-Geisser correction was applied, and the corrected values together with the ϵ correction factor appear in the text where appropriate. η_p^2 effect sizes are also given in the text. The Timing \times Position interaction was entered into a planned comparison. Post-hoc Tukey-HSD tests were employed to further detail other significant results. All parametric statistical analyses were computed using Statistica (version 13.4.0.14; TIBCO Software, Palo Alto, CA, U.S.A.).

In order to check a possible bias in window selection we entered the difference ERPs from all condition-position combinations separately to paired two-tailed t -tests at all time points between 1 and 499 ms at all 9 scalp electrodes (i.e., 4491 total comparisons). The Benjamini and Yekutieli (2001) procedure for control of the false discovery rate (FDR) was applied to assess the significance of each test using an FDR level of 5%. This particular FDR procedure guarantees that the true FDR will be less than or equal to the nominal FDR level of 5% regardless of the dependency structure of the multiple tests. The tests were conducted with the Mass Univariate ERP Toolbox (version 03272017, Gropppe, Urbach, & Kutas, 2011).

In order to quantify the null result of the GLM for the Timing \times Position interaction of the Jittered sequences (e.g., the absence of statistical learning in terms of the hypotheses) we used a Bayesian t -test as implemented in JASP (JASP Team, 2019; Wagenmakers et al., 2018) to compare between the jittered beat and jittered offbeat MMR amplitudes at Cz. We estimated Bayes factors using a Cauchy prior distribution ($r = 0.71$) with no difference between the Beat and Off-beat MMR amplitudes as the null hypothesis. In addition, we performed a robustness check as implemented in JASP to assess whether the results would change with a different prior ($r = 1$) as originally proposed for Bayesian t -tests (Jeffreys, 1961; Wagenmakers et al., 2018).

3. Results

Fig. 3 shows an overview of the central (average of C3, Cz, and C4) responses elicited by deviant and standard stimuli at the Beat and Offbeat positions in the Isochronous and Jittered conditions, together with the corresponding difference waveforms. The difference waveforms for the beat deviants appear as a broad negative wave, which is consistent with the morphology of mismatch response (MMR) in newborn infants (Kushnerenko et al., 2002). Figures including all channels included in the analyses for all grand averages (Supplementary Figs. 1–3) as well as individual averages and 95% confidence intervals (CI) for all grand averages (Supplementary Figs. 4–15) are given in the Supplementary Materials. Descriptive statistics are given in Supplementary Table 2. Fig. 4 shows the distributions of the MMR amplitudes obtained for the Beat and Offbeat positions, separately for the Isochronous and the Jittered conditions.

The Timing \times Position \times Frontality \times Laterality model yielded a significant main effect of Position ($F(1,26) = 5.086$, $p = 0.033$, $\eta_p^2 = 0.164$), which was caused by more negative MMR amplitudes for Beat compared to Offbeat positions ($df = 26$, $p = 0.033$), and a significant main effect of Frontality ($F(1.620, 42.116) = 7.178$, $p = 0.004$, $\epsilon = 0.810$, $\eta_p^2 = 0.216$), caused by less negative MMR amplitudes for frontal compared to both central ($df = 52$, $p = 0.034$) and parietal ($df = 52$, $p =$

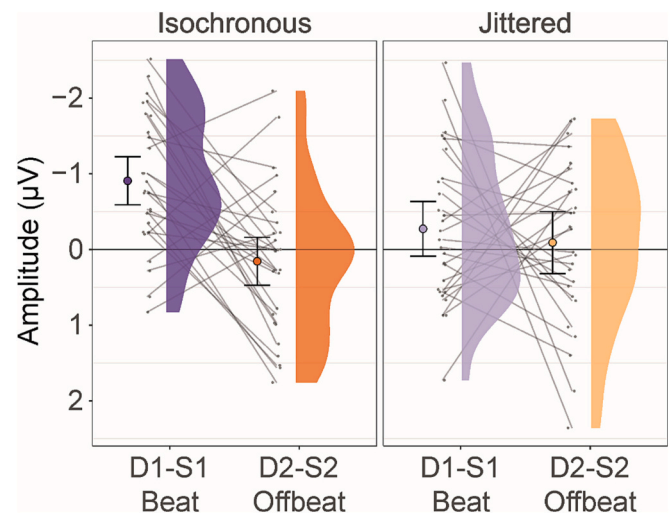


Fig. 4. Distribution of the average MMR amplitudes for all participants and conditions for the average of the three central electrodes (C3, Cz, and C4). Colored dots indicate group average with the error bars denoting 95% confidence intervals. Grey dots connected by lines show the same individual's Beat and Offbeat MMR amplitudes, separately for the two conditions. Note that the labels Beat and Offbeat are used to refer to the position of the deviants in the rhythmic sequence. But a Beat (or Offbeat) can, of course, only be sensed in the isochronous condition, and not in the jittered condition.

0.002) electrodes.

Importantly, there was also a significant Timing \times Position interaction ($F(1, 26) = 5.443$, $p = 0.028$, $\eta_p^2 = 0.173$). Planned comparisons revealed that MMN amplitudes were significantly more negative for Beat responses, compared to Offbeat responses in the Isochronous Timing condition ($F(1) = 13.136$, $p = 0.001$) whereas they did not differ from each other in the Jittered Timing condition ($F(1) = 0.013$, $p = 0.909$) (see Fig. 4). Further, Beat MMR responses were more negative in the Isochronous compared to the Jittered Timing condition ($F(1) = 6.466$, $p = 0.017$), whereas Offbeat responses did not differ between the two conditions ($F(1) = 1.860$, $p = 0.184$). The Bayesian t -test indicated moderate evidence in favor of the null hypothesis (no difference between Beat and Offbeat MMR amplitudes) for the Jittered condition ($BF_{01} = 4.49$). Bayes factors between 3 and 10 are regarded as moderate evidence in favor of the null hypothesis (Wagenmakers et al., 2018). The robustness check indicated that the results did not change as a function of the prior used (with the more traditional prior of $r = 1$: $BF_{01} = 6.135$). The output of the JASP software is included in the Supplementary Materials.

There was also a significant Position \times Laterality effect ($F(1.845, 47.962) = 3.74507576$, $p = 0.034$, $\epsilon = 0.922$, $\eta_p^2 = 0.126$). Post-hoc testing revealed that while MMR amplitudes at the Beat position were significantly more negative compared to the Offbeat position at all electrodes, this difference was smaller on the right side of the scalp ($df = 52$, $p = 0.044$ and $p = 0.034$, respectively, all other differences $p < 0.001$).

The FDR corrected mass univariate t -tests have uncovered significant differences only in the Isochronous condition at the Beat position ($df = 26$, $q < 0.05$). The difference was significant on channels C3, P3, Cz and Pz within the windows (180–348, 182–342, 231–345, 203–355, respectively, all in milliseconds from deviant sound onset). (See Supplementary Figs. 16–19).

4. Discussion

The results show that whereas in the isochronous condition there was a significant difference between the MMR amplitudes obtained for deviants presented at beat and offbeat positions, this difference was absent

in the jittered condition. The dependence of difference responses on metrical position in the isochronous, but not the jittered condition is indicative of the presence of beat processing. These results support and provide converging evidence to the main conclusion of Winkler et al. (2009) that beat detection is already functional at birth in healthy infants. Importantly, the current results show that beat detection in newborns is unlikely to be solely explained by statistical learning, i.e., at least not by the known neonatal ability of learning transitional probabilities of stimulus sequences. Note that here, since we examine newborn responses, we consider beat detection (or perception) to be evident from brain responses even in the absence of overt behavior.

Beat processing (or beat perception) has been argued to be a fundamental component of the capacity for music and a prerequisite for being able to dance and make music together (Honing, ten Cate, Peretz, & Trehub, 2015). Beat processing has been shown to be absent in macaque monkeys when using the same paradigm as in the current study (Honing et al., 2018), but they do sense rhythmic pattern (Honing, Merchant, Håden, Prado, & Bartolo, 2012) and isochrony (Criscuolo et al., 2023). Beat perception (and synchronization) appears to have evolved gradually within the primates (Merchant & Honing, 2014), peaking in humans and present only with limitations in chimpanzees (Hattori & Tomonaga, 2019), bonobos (Large & Gray, 2015), gibbons (Raimondi et al., 2023) and other nonhuman primates (Bouwer, Nityananda, Rouse, & ten Cate, 2021; Honing et al., 2012; Honing et al., 2018), while isochrony perception seems to be shared by most primates (Criscuolo et al., 2023; Honing et al., 2018).

Furthermore, a recent genome-wide association studies (GWAS) provided support for beat based processing exhibiting a highly polygenic architecture, with sixty-seven loci reaching genome-wide significance and a heritability of 13%–16% (Niarchou, Gustavson, Sathirapongsasuti, et al., 2022). As such, beat processing can be considered a biological trait, though it remains unclear what the underlying mechanisms are. The current study is an important first step by providing evidence for beat perception, isochrony perception and statistical learning to be three distinct mechanisms.

When combining the current results with our previous study, we now have converging evidence from two different paradigms suggesting that beat processing is functional in newborn infants (Winkler et al., 2009). This can be taken as support for a biological basis of beat perception, per se (Merchant & Honing, 2014; ten Cate & Honing, 2024). That is, while statistical learning is an essential function for extracting information from the world, the current results suggest that it is complemented by functions honed to extract the temporal structure of (at least) the acoustic environment. Their complementary nature is supported by findings showing that temporal predictability, which is enhanced by isochronous stimulus presentation, appears to improve statistical learning (Selchenkova, Jones, & Tillmann, 2014; Tsogli, Jentschke, & Koelsch, 2022). In fact, much of the research on statistical learning, especially in neonates, has been based on isochronous stimuli (Bosseler et al., 2016; Teinonen et al., 2009). Indeed, it is easy to speculate that extracting the temporal structure of a sound sequence helps to reduce its variability by allowing the brain to focus its processing efforts in time (entrainment, see e.g., Stefanics et al., 2010). A similar assumption has been suggested in the Dynamic Attention Theory (Jones, 1976).

While the main hypothesis of the current study was confirmed by the data, contrary to the previous results in adults (Bouwer et al., 2016), we found no main effect of the temporal schedule of stimulus presentation (isochronous vs. jittered sound delivery). This was due to the lack of difference between offbeat MMN amplitudes in the isochronous and jittered conditions. Temporal predictability, as is present in the isochronous sequences, be it based on its regularity or on the predictable single interval (Bouwer et al., 2021), should enhance processing of events in the isochronous sequence as compared to the jittered sequence (Schwartz, Rothermich, Schmidt-Kassow, & Kotz, 2011). It is unlikely that newborns would be capable of detecting a beat, but not isochrony, which, in effect, is a single level of regularity. Indeed, newborns are

sensitive to temporal aspects of sequences (Håden et al., 2015). More likely, the absence of a difference in the offbeat positions between isochronous and jittered conditions is in fact due to beat perception. Temporal predictability in general should make events in the isochronous condition more expected (better specified), and thus violations of their regularities should elicit larger prediction errors regardless of metrical position. However, one of the main features of beat perception is that on top of enhancing the processing of metrically important events, it also suppresses processing of metrically less important ones (Bouwer et al., 2020; Breska & Deouell, 2017). Possibly, the presence of the beat in the isochronous condition of the current study hampered the processing of the offbeat events, which, in neonates, resulted in no or less specific prediction for offbeat events.

Perhaps the most striking difference between the current study in neonates and previous results with the same paradigm in adults is that in the absence of a clear MMR amplitude difference between beat and offbeat positions in the jittered condition, the current results are inconclusive with regards to the exact role of statistical learning in processing of rhythmic sequences. In the same paradigm, adults have shown different EEG responses to deviants based on position in the jittered sequences even when not attending to the stimuli (Bouwer et al., 2016), although this was significant only for the P3a component. This was argued to reflect the adults' ability to extract statistical regularity from the sequence, even when the timing is not isochronous. Further, statistical learning has been shown to operate in neonates (Bosseler et al., 2016; Bulf et al., 2011; Teinonen et al., 2009). There are several possible explanations of why statistical learning was not observed in the current study. The simplest one is that our measure was not sufficiently sensitive for neonates, whose EEG has a much lower signal-to-noise ratio for ERPs than adults (Kushnerenko, Van den Bergh, & Winkler, 2013). The Bayesian analysis of this null result signified only moderate evidence for no statistical learning, which leaves this question open for future investigations. A limitation on the interpretation of the results is that in adults, MMN amplitude differences between positions in jittered sequences were not shown in the unattended blocks, in which only the P3a showed this effect. It is not entirely clear whether we could expect the MMR to index this effect in sleeping infants. The interpretation is further complicated by the fact that at birth, MMN can vary by stimulus features (Cheng et al., 2015) and whereas the negative MMN shows similarities to the adult MMN (e.g., response to "higher order" rule violations), the positive MMN (P-MMR) shares some features of the adult P3a (e.g., response to contextual novelty, wide-band stimuli). Thus, they cannot be fully identified as infantile precursors of the corresponding adult components (Kushnerenko et al., 2013). In sum, the lack of difference in the response does not rule out statistical learning in jittered sequences in neonates. It only proves that beat processing cannot be explained by statistical learning at birth.

As the current data does not shed further light on this issue, here we only bring up the explanation most compatible with the notion of complementary parallel processes detecting temporal and sequential regularities in the human brain. It is possible that isochrony or, more generally, temporal predictability is important for newborns to extract sequential probabilities necessary for statistical learning. Newborns have no problem extracting even higher order regularities in temporally highly predictable stimuli (e.g., Ruusuvirta, Huotilainen, Fellman, & Näätänen, 2004; Stefanics et al., 2010), and although predictability in time is not necessary for statistical learning in adults, it does enhance the statistical MMN amplitude, an index of statistical learning (Tsogli et al., 2022). Further, in general, the noisier the regularity, the lower the MMN amplitude (e.g., Winkler et al., 1990). In other words, predictions with low specificity result in low-amplitude error signals (Friston, 2005; Southwell & Chait, 2018). We have hypothesized above that extracting temporal structure reduces the observed variability (e.g., by separating the regularities for beat and offbeat sounds) while helping to focus the processing in time (entrainment), substantially reducing the capacities necessary to extract and learn the statistical structure of the input. Thus,

extracting the temporal structure of sound sequences may be crucial for neonates to learn their sequential structure whereas adults can learn this even when the temporal schedule of sound presentation is noisier. Obviously, further studies are needed to test this hypothesis.

The observed Position \times Laterality interaction stemming from larger left-hemisphere MMR effects can be interpreted in terms of general auditory feature processing. Temporal features tend to be processed to a larger extent in the left hemisphere (e.g., Tervaniemi & Hugdahl, 2003) and spectral features to a larger extent on the right side (Zatorre & Belin, 2001). This also appears to be the case in newborn infants (DeCasper & Prescott, 2009). Because processing of temporal features (i.e. isochrony) plays an important role in the current stimulus paradigm and the deviant stimuli do not differ from the standards in spectral features, the observed larger left-hemispheric effects are consistent with the previous literature. On the other hand, much less is known about the lateralization of the processes underlying statistical learning in infants. Although studies using linguistic stimuli typically show left lateralization for statistical learning the results are much less clear for nonlinguistic stimuli which can show either left, right, or even bilateral activation in adults (Moser et al., 2021). Therefore, the current lateralization effect cannot be interpreted in relation to statistical learning.

5. Conclusions

In the current experiment we have established that newborn infants are capable of beat based processing, providing converging evidence for the conclusions of Winkler et al. (2009). Importantly, the paradigm of Bouwer et al. (2016) used here allowed for the separation of beat processing and statistical learning of transition probabilities in neonates. Although the results suggest the presence of beat perception in newborns, we could not show the presence of statistical learning of transition probabilities when sequence timing was not isochronous. Current results, and previous results that show better statistical learning for sequences of regular temporal structure, bring up the possibility that extracting the temporal structure and statistical learning work in a complementary fashion.

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.cognition.2023.105670>.

Ethics

The study was carried out at the Department of Obstetrics-Gynaecology and Perinatal Intensive Care Unit, Military Hospital, Budapest, Hungary. The study was conducted in full accordance with the Declaration of Helsinki and all applicable national laws, and it was approved by the relevant ethics committee: Medical Research Council-Committee of Scientific and Research Ethics (ETT-TUKEB), Hungary.

Funding

This work was supported by the National Research and Innovation Office of Hungary (project no. FK139135 to GPH). GPH was supported by the Bolyai Research Grants from the Hungarian Academy of Sciences (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). FLB was supported by a VENI grant from the Dutch Research Council NWO (VI.Veni.201G.066). HH was supported by an Open Competition grant from the Dutch Research Council NWO (406.20.CW.002).

Author contributions

FLB and HH designed the paradigm, All designed the study, FLB created stimuli, GPH acquired and analyzed data, All interpreted results, GPH wrote manuscript, All edited manuscript, FLB provided additional Bayesian analysis. Both HH and IW are senior authors of this manuscript.

CRediT authorship contribution statement

Gábor P. Hádén: Conceptualization, Investigation, Formal analysis, Writing – original draft, Writing – review & editing. **Fleur L. Bouwer:** Conceptualization, Methodology, Resources, Formal analysis, Writing – review & editing. **Henkjan Honing:** Conceptualization, Methodology, Writing – review & editing, Supervision. **István Winkler:** Conceptualization, Writing – review & editing, Supervision.

Declaration of Competing Interest

The authors declare no conflicts of interest.

Data availability

Preprocessed data, analysis scripts and results will be made publicly available at the OSF repository. (DOI [10.17605/OSF.IO/EFQBK](https://doi.org/10.17605/OSF.IO/EFQBK)).

Acknowledgement

The authors are grateful to all infants and their parents for participating in the experiments, Prof. Dr. Miklós Török and Dr. József Korcsik for supporting the experiments at the Department of Obstetrics, Military Hospital, Budapest, and Judit Roschéné Farkas for recording the data.

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. *Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section*, 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.
- Benjamini, Y., & Yekutieli, D. (2001). The control of the false discovery rate in multiple testing under dependency. *Annals of Statistics*, 1165–1188.
- Bosseler, A. N., Teinonen, T., Tervaniemi, M., & Huotilainen, M. (2016). Infant directed speech enhances statistical learning in newborn infants: An ERP study. *PLoS One*, 11(9).
- Bouwer, F. L., Burgoyne, J. A., Odijk, D., Honing, H., & Grahn, J. A. (2018). What makes a rhythm complex? The influence of musical training and accent type on beat perception. *PLoS ONE*, 13(1).
- Bouwer, F. L., Fahrenfort, J. J., Millard, S. K., Kloosterman, N. A., & Slagter, H. A. (2023). A silent disco: Differential effects of beat-based and pattern-based temporal expectations on persistent entrainment of low-frequency neural oscillations. *Journal of Cognitive Neuroscience*, 35, 990–1020.
- Bouwer, F. L., Honing, H., & Slagter, H. A. (2020). Beat-based and memory-based temporal expectations in rhythm: Similar perceptual effects, different underlying mechanisms. *Journal of Cognitive Neuroscience*, 32, 1221–1241. <https://doi.org/10.1162/jocn.a.01529>
- Bouwer, F. L., Nityananda, V., Rouse, A. A., & ten Cate, C. (2021). Rhythmic abilities in humans and non-human animals: A review and recommendations from a methodological perspective. *Philosophical Transactions of the Royal Society B*, 376, 20200335. <https://doi.org/10.1098/rstb.2020.0335>
- Bouwer, F. L., Van Zuijen, T. L., & Honing, H. (2014). Beat processing is pre-attentive for metrically simple rhythms with clear accents: An ERP study. *PLoS One*, 9(5).
- Bouwer, F. L., Werner, C. M., Knetemann, M., & Honing, H. (2016). Disentangling beat perception from sequential learning and examining the influence of attention and musical abilities on ERP responses to rhythm. *Neuropsychologia*, 85, 80–90.
- Breska, A., & Deouell, L. Y. (2017). Neural mechanisms of rhythm-based temporal prediction: Delta phase-locking reflects temporal predictability but not rhythmic entrainment. *PLoS Biology*, 15(2), Article e2001665.
- Bulf, H., Johnson, S. P., & Valenza, E. (2011). Visual statistical learning in the newborn infant. *Cognition*, 121(1), 127–132.
- Cheng, Y. Y., Wu, H. C., Tzeng, Y. L., Yang, M. T., Zhao, L. L., & Lee, C. Y. (2015). Feature-specific transition from positive mismatch response to mismatch negativity in early infancy: Mismatch responses to vowels and initial consonants. *International Journal of Psychophysiology*, 96(2), 84–94.
- Conway, C. M. (2020). How does the brain learn environmental structure? Ten core principles for understanding the neurocognitive mechanisms of statistical learning. *Neuroscience & Biobehavioral Reviews*, 112, 279–299.
- Cooper, G., & Meyer, L. B. (1963). *The Rhythmic Structure of Music*. University of Chicago press.
- Crisuolo, A., Schwartze, M., Prado, L., Ayala, Y., Merchant, H., & Kotz, S. A. (2023). Macaque monkeys and humans sample temporal regularities in the acoustic environment. *Progress in Neurobiology*, 229, Article 102502.

- DeCasper, A. J., & Prescott, P. (2009). Lateralized processes constrain auditory reinforcement in human newborns. *Hearing Research*, 255(1–2), 135–141.
- Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, 134(1), 9–21.
- Duncan, D., & Theeuwes, J. (2020). Statistical learning in the absence of explicit top-down attention. *Cortex*, 131, 54–65.
- Finlay, B. L., & Uchiyama, R. (2020). The timing of brain maturation, early experience, and the human social niche. In *Evolutionary neuroscience* (pp. 815–843). Academic Press.
- Fiser, J., & Aslin, R. N. (2001). Unsupervised statistical learning of higher-order spatial structures from visual scenes. *Psychological Science*, 12(6), 499–504.
- Fitzgerald, K., & Todd, J. (2020). Making sense of mismatch negativity. *Frontiers in Psychiatry*, 11, 468.
- Friston, K. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 360, 815–836.
- Groppe, D. M., Urbach, T. P., & Kutas, M. (2011). Mass univariate analysis of event-related brain potentials/fields I: A critical tutorial review. *Psychophysiology*, 48(12), 1711–1725.
- Håden, G. P., Honing, H., Török, M., & Winkler, I. (2015). Detecting the temporal structure of sound sequences in newborn infants. *International Journal of Psychophysiology*, 96(1), 23–28.
- Håden, G. P., Németh, R., Török, M., & Winkler, I. (2016). Mismatch response (MMR) in neonates: Beyond refractoriness. *Biological Psychology*, 117, 26–31.
- Hattori, Y., & Tomonaga, M. (2019). Rhythmic swaying induced by sound in chimpanzees (*Pan troglodytes*). *Proceedings of the National Academy of Sciences*, 117(2), 936–942.
- Honing, H. (2012). Without it no music: Beat induction as a fundamental musical trait. *Annals of the New York Academy of Sciences*, 1252(1), 85–91.
- Honing, H., & Bouwer, F. L. (2019). Rhythm. In P. J. Rentfrow, & D. Levitin (Eds.), *Foundations in music psychology: Theory and research*. Cambridge, Mass: The MIT Press (ISBN 9780262039277.).
- Honing, H., Bouwer, F. L., Prado, L., & Merchant, H. (2018). Rhesus monkeys (*Macaca mulatta*) sense isochrony in rhythm, but not the beat: Additional support for the gradual audiomotor evolution hypothesis. *Frontiers in Neuroscience*, 12, 475.
- Honing, H., & Merchant, H. (2014). Differences in auditory timing between human and nonhuman primates. *Behavioral and Brain Sciences*, 37(06), 557–558.
- Honing, H., Merchant, H., Håden, G. P., Prado, L., & Bartolo, R. (2012). Rhesus monkeys (*Macaca mulatta*) detect rhythmic groups in music, but not the beat. *PLoS One*, 7(12), 1–10.
- Honing, H., ten Cate, C., Peretz, I., & Trehub, S. E. (2015). Without it no music: Cognition, biology and evolution of musicality. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 370(1664), 20140088. <https://doi.org/10.1098/rstb.2014.0088>
- JASP Team. (2019). JASP (version 0.9.2) [computer software]. Retrieved from <http://jasp-stats.org>.
- Jeffreys, H. (1961). *Theory of probability* (3rd ed.). Oxford: Oxford University Press.
- Jones, M. R. (1976). Time, our lost dimension: Toward a new theory of perception, attention, and memory. *Psychological Review*, 83, 323–355.
- Kotz, S. A., Ravignani, A., & Fitch, W. T. (2018). The evolution of rhythm processing. *Trends in Cognitive Sciences*, 22(10), 896–910.
- Kushnerenko, E., Ceponiene, R., Balan, P., Fellman, V., Huotilainen, M., & Näätänen, R. (2002). Maturation of the auditory event-related potentials during the first year of life. *Neuroreport*, 13(1), 47–51.
- 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.
- Ladinig, O., Honing, H., Håden, G., & Winkler, I. (2009). Probing attentive and preattentive emergent meter in adult listeners without extensive music training. *Music Perception: An Interdisciplinary Journal*, 26(4), 377–386.
- Langus, A., Mehler, J., & Nespore, M. (2017). Rhythm in language acquisition. *Neuroscience & Biobehavioral Reviews*, 81, 158–166.
- Large, E. W., & Gray, P. M. (2015). Spontaneous tempo and rhythmic entrainment in a bonobo (*Pan paniscus*). *Journal of Comparative Psychology*, 129(4), 317.
- Large, E. W., Herrera, J. A., & Velasco, M. J. (2015). Neural networks for beat perception in musical rhythm. *Frontiers in Systems Neuroscience*, 9, 159.
- London, J. (2002). Cognitive constraints on metric systems: Some observations and hypotheses. *Music Perception: An Interdisciplinary Journal*, 19(4), 529–550.
- Lopez-Calderon, J., & Luck, S. J. (2014). ERPLAB: An open-source toolbox for the analysis of event-related potentials. *Frontiers in Human Neuroscience*, 8, 213.
- Merchant, H., & Honing, H. (2014). Are non-human primates capable of rhythmic entrainment? Evidence for the gradual audiomotor evolution hypothesis. *Frontiers in Neuroscience*, 7(8 JAN), 1–8.
- Moser, J., Batterink, L., Hegner, Y. L., Schlegel, F., Braun, C., Paller, K. A., & Preissl, H. (2021). Dynamics of nonlinguistic statistical learning: From neural entrainment to the emergence of explicit knowledge. *Neuroimage*, 240, Article 118378.
- Niarchou, M., Gustavson, D. E., Sathirapongsasuti, J. F., et al. (2022). Genome-wide association study of musical beat synchronization demonstrates high polygenicity. *Nature Human Behaviour*, 6, 1292–1309.
- Novitski, N., Tervaniemi, M., Huotilainen, M., & Näätänen, R. (2004). Frequency discrimination at different frequency levels as indexed by electrophysiological and behavioral measures. *Cognitive Brain Research*, 20(1), 26–36.
- Patel, A. D. (2008). *Music, Language, and the Brain*. Oxford: Oxford University Press.
- Pearce, M. T. (2018). Statistical learning and probabilistic prediction in music cognition: Mechanisms of stylistic enculturation. *Annals of the New York Academy of Sciences*, 1423(1), 378–395.
- Polich, J. (2007). Updating P300: An integrative theory of P3a and P3b. *Clinical Neurophysiology*, 118(10), 2128–2148.
- Raimondi, T., Di Panfilo, G., Pasquali, M., Zarantonello, M., Favaro, L., Savini, T., Gamba, M., & Ravignani, A. (2023). Isochrony and rhythmic interaction in aperiodic duetting. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 290, 20222244.
- Raviv, L., & Arnon, I. (2018). The developmental trajectory of children's auditory and visual statistical learning abilities: Modality-based differences in the effect of age. *Developmental Science*, 21(4), Article e12593.
- Ruusuvirta, T., Huotilainen, M., Fellman, V., & Näätänen, R. (2004). Newborn human brain identifies repeated auditory feature conjunctions of low sequential probability. *European Journal of Neuroscience*, 20(10), 2819–2821.
- 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.
- Saffran, J. R., Newport, E. L., & Aslin, R. N. (1996). Word segmentation: The role of distributional cues. *Journal of Memory and Language*, 35(4), 606–621.
- Saffran, J. R., Newport, E. L., Aslin, R. N., Tunick, R. A., & Barrueco, S. (1997). Incidental language learning: Listening (and learning) out of the corner of your ear. *Psychological Science*, 8(2), 101–105.
- Schwartz, M., Rothermich, K., Schmidt-Kassow, M., & Kotz, S. A. (2011). Temporal regularity effects on pre-attentive and attentive processing of deviance. *Biological Psychology*, 87(1), 146–151.
- Selchenkova, T., Jones, M. R., & Tillmann, B. (2014). The influence of temporal regularities on the implicit learning of pitch structures. *Quarterly Journal of Experimental Psychology*, 67(12), 2360–2380.
- Southwell, R., & Chait, M. (2018). Enhanced deviant responses in patterned relative to random sound sequences. *Cortex*, 109, 92–103.
- Stefanics, G., Hangya, B., Hernádi, I., Winkler, I., Lakatos, P., & Ulbert, I. (2010). Phase entrainment of human delta oscillations can mediate the effects of expectation on reaction speed. *The Journal of Neuroscience*, 30(41), 13578–13585.
- Teinonen, T., Fellman, V., Näätänen, R., Alku, P., & Huotilainen, M. (2009). Statistical language learning in neonates revealed by event-related brain potentials. *BMC Neuroscience*, 10(1), 21.
- Teki, S., Grube, M., Kumar, S., & Griffiths, T. D. (2011). Distinct neural substrates of duration-based and beat-based auditory timing. *The Journal of Neuroscience*, 31(10), 3805–3812.
- ten Cate, C., & Honing, H. (2024). Precursors of music and language in animals. In D. Sammler (Ed.), *Oxford handbook of language and music*. Oxford: Oxford University Press. Preprint: psyarxiv.com/4zxtr.
- Tervaniemi, M., & Hugdahl, K. (2003). Lateralization of auditory-cortex functions. *Brain Research Reviews*, 43(3), 231–246.
- Tsoglis, V., Jentschke, S., & Koelsch, S. (2022). Unpredictability of the “when” influences prediction error processing of the “what” and “where”. *PLoS One*, 17(2), Article e0263373.
- Wagenmakers, E. J., Love, J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., ... Morey, R. D. (2018). Bayesian inference for psychology. Part II: Example applications with JASP. *Psychonomic Bulletin & Review*, 25(1), 58–76.
- Winkler, I., Håden, G. P., Ladinig, O., Sziller, I., & Honing, H. (2009). Newborn infants detect the beat in music. *Proceedings of the National Academy of Sciences*, 106(7), 2468–2471.
- Winkler, I., Paavilainen, P., Alho, K., Reinikainen, K., Sams, M., & Näätänen, R. (1990). The effect of small variation of the frequent auditory stimulus on the event related brain potential to the infrequent stimulus. *Psychophysiology*, 27, 228–235.
- Zatorre, R. J., & Belin, P. (2001). Spectral and temporal processing in human auditory cortex. *Cerebral Cortex*, 11(10), 946–953.