bioRxiv preprint doi: https://doi.org/10.1101/2022.12.20.521245; this version posted December 21, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

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

Gábor P. Háden[1,2]*, Fleur L. Bouwer[3,4,5], Henkjan Honing[3,4], István Winkler[1]

[1]Institute of Cognitive Neuroscience and Psychology, Research Centre for Natural Sciences, H-1117 Budapest, Magyar tudósok körútja 2., Hungary

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

[3] Music Cognition Group, Institute for Logic, Language, and Computation, University of Amsterdam, P.O. Box 94242, 1090 GE Amsterdam, The Netherlands

[4] Amsterdam Brain and Cognition, University of Amsterdam, P.O. Box 15900, 1001 NK Amsterdam, The Netherlands

[5] Department of Psychology, Brain & Cognition, University of Amsterdam, P.O. Box 15900, 1001 NK Amsterdam, The Netherlands

*Corresponding author.

Gábor P. Háden, Institute of Cognitive Neuroscience and Psychology, Research Centre for Natural Sciences, Magyar tudósok körútja 2., Budapest H-1117, Hungary. Telephone: +36 1 382 6808, Fax: +36 1 3826295, ORCID: 0000-0002-7664-9653 Email: haden.gabor@ttk.hu

Author details:
Fleur Bouwer, ORCID: 0000-0001-7064-7481, email: f.l.bouwer@uva.nl
Henkjan Honing, ORCID: 0000-0001-7046-7923, email: honing@uva.nl
István Winkler, ORCID: 0000-0002-3344-6151, email: winkler.istvan@ttk.hu

Data Statement
Preprocessed data, analysis scripts and results will be made publicly available at the OSF repository. (DOI 10.17605/OSF.IO/EFQBK)

Declarations of interest
The authors declare no conflicts of interest.

Funding
This work was supported by the National Research and Innovation Office of Hungary (project no. K115385 to IW). 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-21-5-BME-364). 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).

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.

bioRxiv preprint doi: https://doi.org/10.1101/2022.12.20.521245; this version posted December 21, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

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.

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.

bioRxiv preprint doi: https://doi.org/10.1101/2022.12.20.521245; this version posted December 21, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Research highlights

Sleeping newborns process musical beat.

Transition probabilities are not enough to explain beat perception in newborn infants.

No evidence of statistical learning (based on transition probabilities) without isochronous stimulation in newborns.

Results converge with previous evidence on beat perception of newborn infants.

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 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. However, the current paradigm did not show effects of statistical learning, despite previous evidence for this ability in newborns. These results show that statistical learning does not explain beat processing in newborn infants.

Keywords: neonate; EEG; rhythm, beat perception, statistical learning

bioRxiv preprint doi: https://doi.org/10.1101/2022.12.20.521245; this version posted December 21, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Introduction

Processing sound sequences is crucial for both music and speech perception (Patel, 2008; Kotz et al., 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, & Nespor, 2017). While both the perception of temporal (Winkler et al., 2009) and statistical (Bosseler et al., 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 et al., 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 (Honing & Bouwer, 2019; Honing, 2012; Cooper & Meyer, 1960). 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 & Valesco, 2015). Beat-based timing has been suggested to be somewhat separate from other timing processes, like the perception of (a sequence of) single absolute temporal intervals (Honing & Merchant, 2014; Teki et al., 2011; Breska & Deouell, 2017; Bouwer, Honing & Slagter, 2020). In humans, we have previously suggested that beat

bioRxiv preprint doi: https://doi.org/10.1101/2022.12.20.521245; this version posted December 21, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

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 (Bouwer et al., 2014). Hence, the presence of beat perception at birth remains to be confirmed. (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 infants cannot describe their experience, making this term incorrect in the literal sense.)

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, et al., 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, Newport & Aslin, 1996), music (Saffran, et al., 1999, Pearce, 2018), and vision (Fiser & Aslin, 2001; Duncan & Theeuwes, 2020). There is evidence that statistical learning is already functional at birth (Bulf, Johnson & Valenza, 2011; Bosseler et al., 2016, Teinonen, et al., 2009). However, there is an ongoing debate whether statistical learning in children is an unchanging domain general mechanism or that it shows developmental changes that differ between the visual and the auditory modality (e.g., Saffran et al., 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-based perception (Bouwer et al., 2016, 2020; Honing et al., 2018). Secondly, while the presence of acoustic variation in a

bioRxiv preprint doi: https://doi.org/10.1101/2022.12.20.521245; this version posted December 21, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

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 statistical learning.

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 (Figure 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 et al., 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 (Figure 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, 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

bioRxiv preprint doi: https://doi.org/10.1101/2022.12.20.521245; this version posted December 21, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

sounds with different timbres.

Bouwer et al. (2016) presented these sequences to adults, and measured event-related electric brain responses (ERPs) to rare, unexpected intensity decrements (i.e., deviants) to assess the formation of beat-based expectations. If 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 et al., 2004). Bouwer et al. (2016) found larger MMN responses for deviants in odd (beat) than even (offbeat) positions, even in the jittered condition, indicative of statistical learning that resulted in different predictions for beat and offbeat positions. The difference was, however, much larger in the isochronous than in the jittered condition, which the authors took as support for the presence of beat perception. Arguably, beat perception contributed to differences between positions only in the isochronous, but not the jittered condition. Moreover, in both 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.

A mismatch response (MMR) similar to MMN can be measured in young infants in response to regularity violations, including sleeping neonates (Alho et al., 1990, Háden et al., 2016). In our previous study in newborn infants (Winkler et al., 2009) stimulus omissions (Háden et al., 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

bioRxiv preprint doi: https://doi.org/10.1101/2022.12.20.521245; this version posted December 21, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

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.'s (2016) stimulus paradigm to sleeping newborn infants to separate the effects of beat perception from those of statistical learning (Figure 1) in order to provide converging evidence to the notion that newborns can perceive a 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.'s (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.

[add Figure 1 around here]

Material and methods

bioRxiv preprint doi: https://doi.org/10.1101/2022.12.20.521245; this version posted December 21, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

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 less than 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.

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 Figure 1 for a schematic depiction of the rhythmic stimulus patterns used, Figure 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

bioRxiv preprint doi: https://doi.org/10.1101/2022.12.20.521245; this version posted December 21, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

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 uneven 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 (London, 2002; Honing, 2012) (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 225ms. 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

bioRxiv preprint doi: https://doi.org/10.1101/2022.12.20.521245; this version posted December 21, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

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 Figure 2. This transition network shows the transition probabilities between the three individual sounds used in both the isochronous and jittered conditions.

[add Figure 2 around here]

Procedure

Stimuli were presented in 6 blocks with 3 blocks consisting of isochronous and 3 blocks of jittered sequences (1,300 patterns, i.e., 2,600 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.)

bioRxiv preprint doi: https://doi.org/10.1101/2022.12.20.521245; this version posted December 21, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

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 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 1.5–30 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 μV 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 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

bioRxiv preprint doi: https://doi.org/10.1101/2022.12.20.521245; this version posted December 21, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

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] × Position [Beat vs. Offbeat] × Frontality [F vs. C vs. P electrode line] × 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 × 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 quantify the null result of the GLM for the Timing × 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 = .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 (Wagenmakers et al., 2018; Jeffreys, 1961).

Results

Figure 3 shows an overview of the central (average of C3, Cz, and C4) responses elicited by deviant and standard stimuli at the Beat and Off-beat 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

bioRxiv preprint doi: https://doi.org/10.1101/2022.12.20.521245; this version posted December 21, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

channels included in the analyses for all grand averages (Supplementary Figures 1-3) as well as individual averages and 95% confidence intervals (CI) for all grand averages (Supplementary Figures 4-15) are given in the Supplementary Materials. Descriptive statistics are given in Supplementary Table 2. Figure 4 shows the distributions of the MMR amplitudes obtained for the Beat and Off-beat positions, separately for the Isochronous and the Jittered conditions.

[add Figure 3 around here]

[add Figure 4 around here]

The Timing × Position × Frontality × 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 ($F(1.620, 42.116)= 7.178$, $p= 0.004$, $\varepsilon= 0.810$, $\eta_p^2= 0.216$), caused by less negative MMR amplitudes for frontal compared to both central (df=52, p=0.034) and parietal (df=52, p=0.002) electrodes.

Importantly, there was also a significant Timing × 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 Figure 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

bioRxiv preprint doi: https://doi.org/10.1101/2022.12.20.521245; this version posted December 21, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

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 × Laterality effect (F(1.845, 47.962)= 3.74507576, p= 0.034, ε= 0.922, η_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, the difference was smaller between the right scalp location of the Beat position and the right and central scalp locations of the Offbeat position (df=52, p=0.044 and p= 0.034, respectively, all other differences p<0.001).

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 and colleagues (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.

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 et al., 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). Beat perception (and synchronization) appears to have evolved gradually within the primates (Merchant & Honing,

bioRxiv preprint doi: https://doi.org/10.1101/2022.12.20.521245; this version posted December 21, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

2014), peaking in humans and present only with limitations in chimpanzees (Hattori & Tomonaga, 2019), bonobos (Large & Gray, 2015), and other nonhuman primates (Honing et al., 2012, Honing et al., 2018, Bouwer et al., 2021).

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, in press). 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 (Tsogli, Jentschke & Koelsch, 2022, Selchenkova, Jones & Tillmann, 2014). 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 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 MMR 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, Nityananda, Rouse & ten Cate, 2021), should enhance processing of events in the isochronous sequence as compared to the jittered sequence (Schwartze

bioRxiv preprint doi: https://doi.org/10.1101/2022.12.20.521245; this version posted December 21, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

et al., 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 (Háden, Honing, Török, & Winkler, 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). This was argued to reflect the adults' ability to extract statistical regularity from the sequence, even when the timing is irregular. Further, statistical learning has been shown to operate in neonates (Bulf, Johnson & Valenza, 2011; Bosseler et al., 2016, 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 &

bioRxiv preprint doi: https://doi.org/10.1101/2022.12.20.521245; this version posted December 21, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

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.

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 et al., 2004; Stefanics et al., 2009), 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, Jentschke & Koelsch, 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). If, as we have hypothesized above, 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), it may substantially reduce the necessary capacities, which may be even more crucial for neonates than for adults. Obviously, further studies are needed to test this hypothesis.

Some indirect support for this hypothesis may be gained from the observed significant Position × Laterality interaction. The processing of different auditory features is thought to be lateralized (e.g., Tervaniemi & Hugdahl, 2003): left for temporal features and right for spectral features (Zatorre & Belin, 2001), which seems to be the case also for newborns (DeCasper & Prescott, 2009). Entrainment to a structured sequence in a statistical learning task showed similar lateralization of brain activity to the current one compared to a non-structured sequence in adults with better performance (Moser et al., 2021). This brings up the possibility of an interaction

bioRxiv preprint doi: https://doi.org/10.1101/2022.12.20.521245; this version posted December 21, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

between entrainment and statistical learning, as we have proposed above.

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 irregular. 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.

bioRxiv preprint doi: https://doi.org/10.1101/2022.12.20.521245; this version posted December 21, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

References

Abboub, N., Nazzi, T., & Gervain, J. (2016). Prosodic grouping at birth. *Brain and language*, *162*, 46-59.

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.

Aslin, R. N., & Newport, E. L. (2012). Statistical learning: From acquiring specific items to forming general rules. *Current directions in psychological science*, *21*(**3**), 170-176.

Barnet, A. B., Ohlrich, E. S., Weiss, I. P., & Shanks, B. (1975). Auditory evoked potentials during sleep in normal children from ten days to three years of age. *Electroencephalography and clinical neurophysiology*, *39*(**1**), 29-41.

Batterink, L. J., & Zhang, S. (2022). Simple statistical regularities presented during sleep are detected but not retained. *Neuropsychologia*, *164*, 108106.

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., 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.

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. doi: 10.1098/rstb.2020.0335

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. doi: 10.1162/jocn_a_01529.

Bulf, H., Johnson, S. P., & Valenza, E. (2011). Visual statistical learning in the newborn infant. *Cognition*, *121*(**1**), 127-132.

bioRxiv preprint doi: https://doi.org/10.1101/2022.12.20.521245; this version posted December 21, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

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.

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.

Draganova, R., Eswaran, H., Murphy, P., Huotilainen, M., Lowery, C., Preissl, H. (2005) Sound frequency change detection in fetuses and newborns, a magnetoencephalographic study. *NeuroImage*, 28, 354–361.

Draganova, R., Eswaran, H., Murphy, P., Lowery, C., Preissl, H. (2007) Serial magnetoencephalographic study of fetal and newborn auditory discriminative evoked responses. *Early Human Development*, 83, 199–207.

Duncan, D., & Theeuwes, J. (2020). Statistical learning in the absence of explicit top-down attention. *Cortex*, *131*, 54-65.

Elazar, A., Alhama, R. G., Bogaerts, L., Siegelman, N., Baus, C., & Frost, R. (2022). When the "Tabula" is Anything but "Rasa:" What Determines Performance in the Auditory Statistical Learning Task?. *Cognitive Science*, *46*(**2**), e13102.

Emberson, L. L., Conway, C. M., & Christiansen, M. H. (2011). Timing is everything: Changes in presentation rate have opposite effects on auditory and visual implicit statistical learning. *Quarterly Journal of Experimental Psychology*, *64*(**5**), 1021-1040.

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.

François, C., Teixidó, M., Takerkart, S., Agut, T., Bosch, L., & Rodriguez-Fornells, A. (2017). Enhanced neonatal brain responses to sung streams predict vocabulary outcomes by age 18 months. *Scientific reports*, *7*(**1**), 1-13.

Friederici, A. D., Mueller, J. L., & Oberecker, R. (2011). Precursors to natural grammar learning: preliminary evidence from 4-month-old infants. *PLoS One*, *6*(**3**), e17920.

bioRxiv preprint doi: https://doi.org/10.1101/2022.12.20.521245; this version posted December 21, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Friston, K. (2005). A theory of cortical responses. Philosophical Transactions of the Royal Society of London, *Series B Biological Sciences*, 360, 815–836

Gervain, J., Macagno, F., Cogoi, S., Peña, M., & Mehler, J. (2008). The neonate brain detects speech structure. *Proceedings of the National Academy of Sciences*, *105*(37), 14222-14227.

Háden, G. P., Németh, R., Török, M., & Winkler, I. (2016). Mismatch response (MMR) in neonates: Beyond refractoriness. *Biological psychology*, *117*, 26-31.

Háden, 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.

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.

Hay, J. F., & Saffran, J. R. (2012). Rhythmic grouping biases constrain infant statistical learning. *Infancy*, *17*(**6**), 610-641.

Henry, M. J., & Herrmann, B. (2014). Low-frequency neural oscillations support dynamic attending in temporal context. *Timing & Time Perception*, *2*(**1**), 62-86.

Hepper, P. G., & Shahidullah, B. S. (1994). The development of fetal hearing. *Fetal and Maternal Medicine Review*, *6*(**3**), 167-179.

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., 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

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., & Bouwer, F. L. (2019). Rhythm. In Rentfrow, P.J., & Levitin, D. (ed.), Foundations in Music Psychology: Theory and Research. Cambridge, Mass.: The MIT Press. ISBN 9780262039277.

Honing, H., Merchant, H., Háden, 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.

JASP Team. (2019). JASP (version 0.9.2) [computer software]. Retrieved from https://jasp-stats.org

bioRxiv preprint doi: https://doi.org/10.1101/2022.12.20.521245; this version posted December 21, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

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 Reviews, 83, 323–355.

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

Kotilahti, K., Nissilä, I., Näsi, T., Lipiäinen, L., Noponen, T., Meriläinen, P., ... & Fellman, V. (2010). Hemodynamic responses to speech and music in newborn infants. *Human brain mapping*, *31*(**4**), 595-603.

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.

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.

Ladinig, O., Honing, H., Háden, 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.

Ladinig O., Honing H., Háden G. P., Winkler I. (2011) Erratum to: Probing attentive and pre-attentive emergent meter in adult listeners with no extensive music training. *Music Perception: An Interdisciplinary Journal*, *26*, 444.

Langus, A., Mehler, J., & Nespor, M. (2017). Rhythm in language acquisition. *Neuroscience & Biobehavioral Reviews*, *81*, 158-166.

Large, E. W., & Jones, M. R. (1999). The dynamics of attending: how people track time-varying events. *Psychological review*, *106*(**1**), 119.

Large, E., & Snyder, J. (2009). Pulse and meter as neural resonance. *Annals of the New York Academy of Sciences*, *1169*(**1**), 46-57.

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.

Leppänen, 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*, 91-101.

bioRxiv preprint doi: https://doi.org/10.1101/2022.12.20.521245; this version posted December 21, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Lerdahl, F., & Jackendoff, R. (1983). An overview of hierarchical structure in music. *Music Perception*, *1*(**2**), 229-252.

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.

Lumaca, M., Trusbak Haumann, N., Brattico, E., Grube, M., & Vuust, P. (2019). Weighting of neural prediction error by rhythmic complexity: A predictive coding account using mismatch negativity. *European Journal of Neuroscience*, *49*(**12**), 1597-1609.

Marcovitch, S., & Lewkowicz, D. J. (2009). Sequence learning in infancy: The independent contributions of conditional probability and pair frequency information. *Developmental Science*, *12*(**6**), 1020-1025.

Martins, M. D., Gingras, B., Puig-Waldmueller, E., & Fitch, W. T. (2017). Cognitive representation of "musical fractals": Processing hierarchy and recursion in the auditory domain. *Cognition*, *161*, 31-45.

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., Schleger, 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*, 118378.

Näätänen, R., Paavilainen, P., Rinne, T., & Alho, K. (2007). The mismatch negativity (MMN) in basic research of central auditory processing: a review. *Clinical neurophysiology*, *118*(12), 2544-2590.

Näätänen, R., Tervaniemi, M., Sussman, E., Paavilainen, P., & Winkler, I. (2001). 'Primitive intelligence'in the auditory cortex. *Trends in neurosciences*, *24*(**5**), 283-288.

Nazzi, T., & Ramus, F. (2003). Perception and acquisition of linguistic rhythm by infants. *Speech communication*, *41*(**1**), 233-243.

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.

Paavilainen, P. (2013). The mismatch-negativity (MMN) component of the auditory event-related potential to violations of abstract regularities: a review. *International journal of psychophysiology*, *88*(**2**), 109-123.

bioRxiv preprint doi: https://doi.org/10.1101/2022.12.20.521245; this version posted December 21, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

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.

Perani, D., Saccuman, M. C., Scifo, P., Spada, D., Andreolli, G., Rovelli, R., ... & Koelsch, S. (2010). Functional specializations for music processing in the human newborn brain. *Proceedings of the National Academy of Sciences*, *107*(**10**), 4758-4763.

Polich, J., 2007. Updating P300: an integrative theory of P3a and P3b. Clinical Neurophysiology 118 (**10**), 2128–2148.

Portas, C. M., Krakow, K., Allen, P., Josephs, O., Armony, J. L., & Frith, C. D. (2000). Auditory processing across the sleep-wake cycle: simultaneous EEG and fMRI monitoring in humans. *Neuron*, *28*(**3**), 991-999.

Provasi, J., Anderson, D. I., & Barbu-Roth, M. (2014). Rhythm perception, production, and synchronization during the perinatal period. *Frontiers in Psychology*, *5*, 1048.

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**), e12593.

Rosenthal, M. A. (2016). Hemispheric asymmetry in the hierarchical perception of music and speech. *Psychological Bulletin*, *142*(**11**), 1165.

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., 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.

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., & Kirkham, N. Z. (2018). Infant statistical learning. *Annual Review of Psychology*, *69*.

Schapiro, A., & Turk-Browne, N. (2015). Statistical learning. In: Arthur W. Toga Ed. Brain Mapping: An Encyclopedic Reference, vol. 3: 501-506. Amsterdam: Elsevier.

bioRxiv preprint doi: https://doi.org/10.1101/2022.12.20.521245; this version posted December 21, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

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.

Slone, L. K., & Johnson, S. P. (2015). Infants' statistical learning: 2-and 5-month-olds' segmentation of continuous visual sequences. *Journal of experimental child psychology*, *133*, 47-56.

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.

Suppanen, E., Huotilainen, M., & Ylinen, S. (2019). Rhythmic structure facilitates learning from auditory input in newborn infants. *Infant Behavior and Development*, *57*, 101346.

Taga, G., Watanabe, H., & Homae, F. (2018). Developmental changes in cortical sensory processing during wakefulness and sleep. *NeuroImage*, *178*, 519-530.

Tal, I., Large, E. W., Rabinovitch, E., Wei, Y., Schroeder, C. E., Poeppel, D., & Golumbic, E. Z. (2017). Neural entrainment to the beat: The "missing-pulse" phenomenon. *Journal of Neuroscience*, *37*(**26**), 6331-6341.

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.

ten Cate, C. & Honing.H. (in press). Precursors of music and language in animals. In Sammler, D. (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.

Tsogli, V., Jentschke, S., & Koelsch, S. (2022). Unpredictability of the "when" influences prediction error processing of the "what" and "where". *PloS one*, *17*(**2**), e0263373.

Van Hedger, S. C., Johnsrude, I. S., & Batterink, L. J. (2022). Musical instrument familiarity affects statistical learning of tone sequences. *Cognition*, *218*, 104949.

Vannasing, P., Florea, O., González-Frankenberger, B., Tremblay, J., Paquette, N., Safi, D., ... & Gallagher, A. (2016). Distinct hemispheric specializations for native and non-native languages in one-day-old newborns identified by fNIRS. *Neuropsychologia*, *84*, 63-69.

Wagenmakers, E. J., Love, J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., ... & Morey,

bioRxiv preprint doi: https://doi.org/10.1101/2022.12.20.521245; this version posted December 21, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

R. D. (2018). Bayesian inference for psychology. Part II: Example applications with JASP. *Psychonomic bulletin & review*, *25*(**1**), 58-76.

Winkler, I. (2007). Interpreting the mismatch negativity. *Journal of Psychophysiology*, *21*(**3-4**), 147.

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.

Winkler, I., Háden, 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.

Zatorre, R. J., & Belin, P. (2001). Spectral and temporal processing in human auditory cortex. *Cerebral cortex*, *11*(**10**), 946-953.

bioRxiv preprint doi: https://doi.org/10.1101/2022.12.20.521245; this version posted December 21, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Figure captions

Figure 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, these are only perceptually valid in the isochronous condition. (Adapted from Honing et al., 2018).

Figure 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 Honing et al., 2018).

Figure 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

bioRxiv preprint doi: https://doi.org/10.1101/2022.12.20.521245; this version posted December 21, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

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.

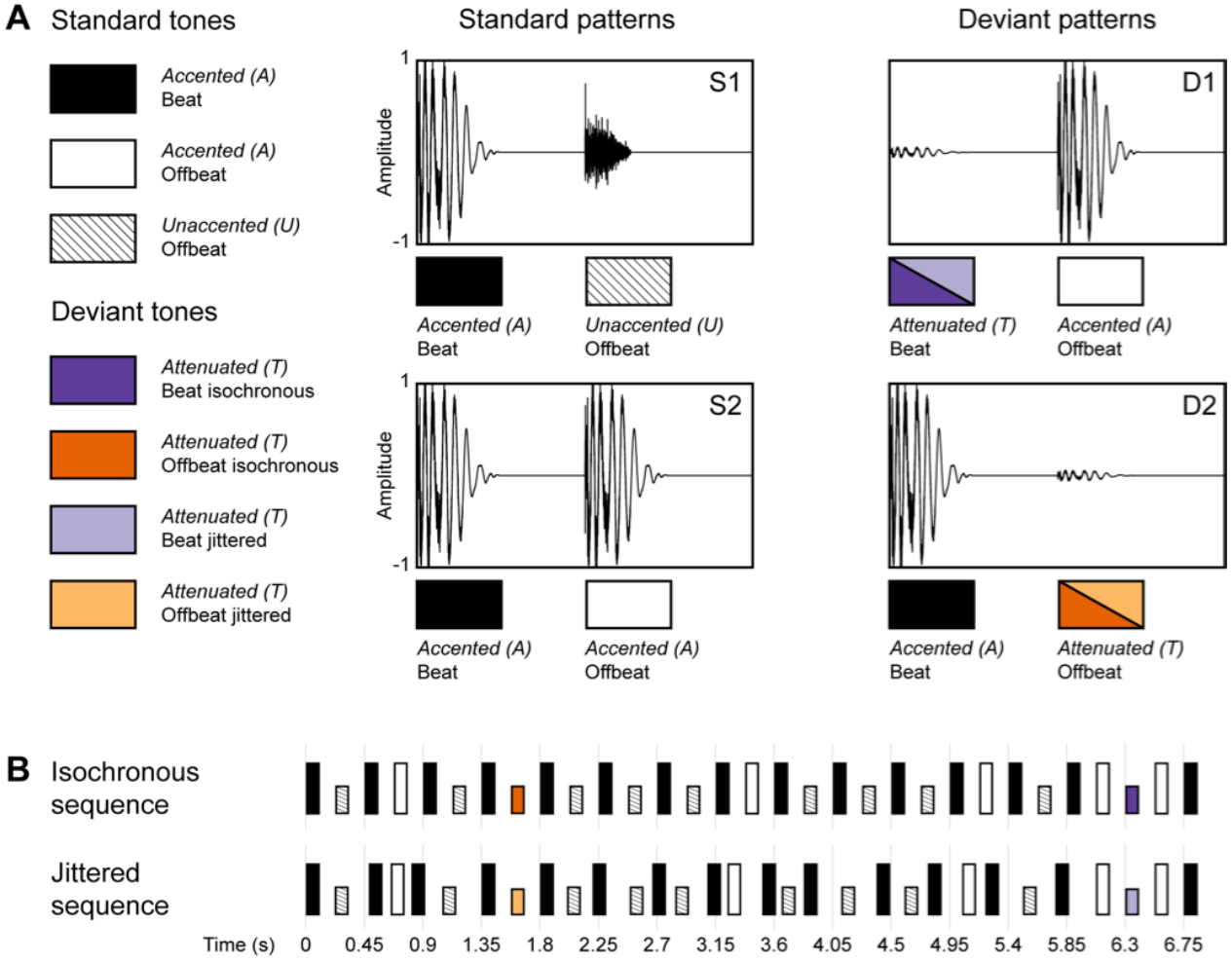
Figure 4

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 Off-beat MMR amplitudes, separately for the two conditions.

bioRxiv preprint doi: https://doi.org/10.1101/2022.12.20.521245; this version posted December 21, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

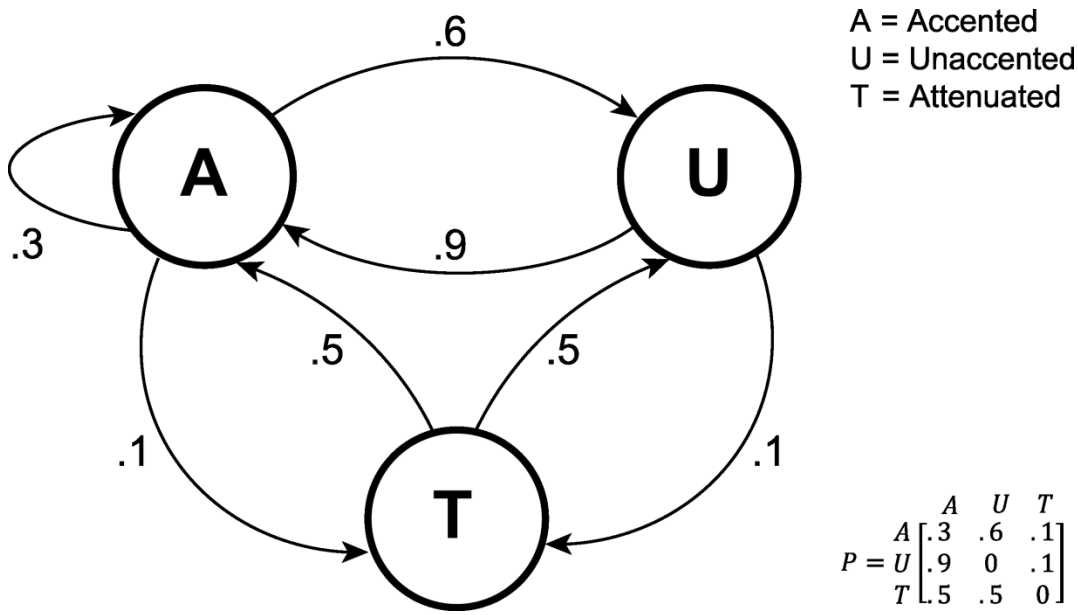
Figures

Figure 1



bioRxiv preprint doi: https://doi.org/10.1101/2022.12.20.521245; this version posted December 21, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

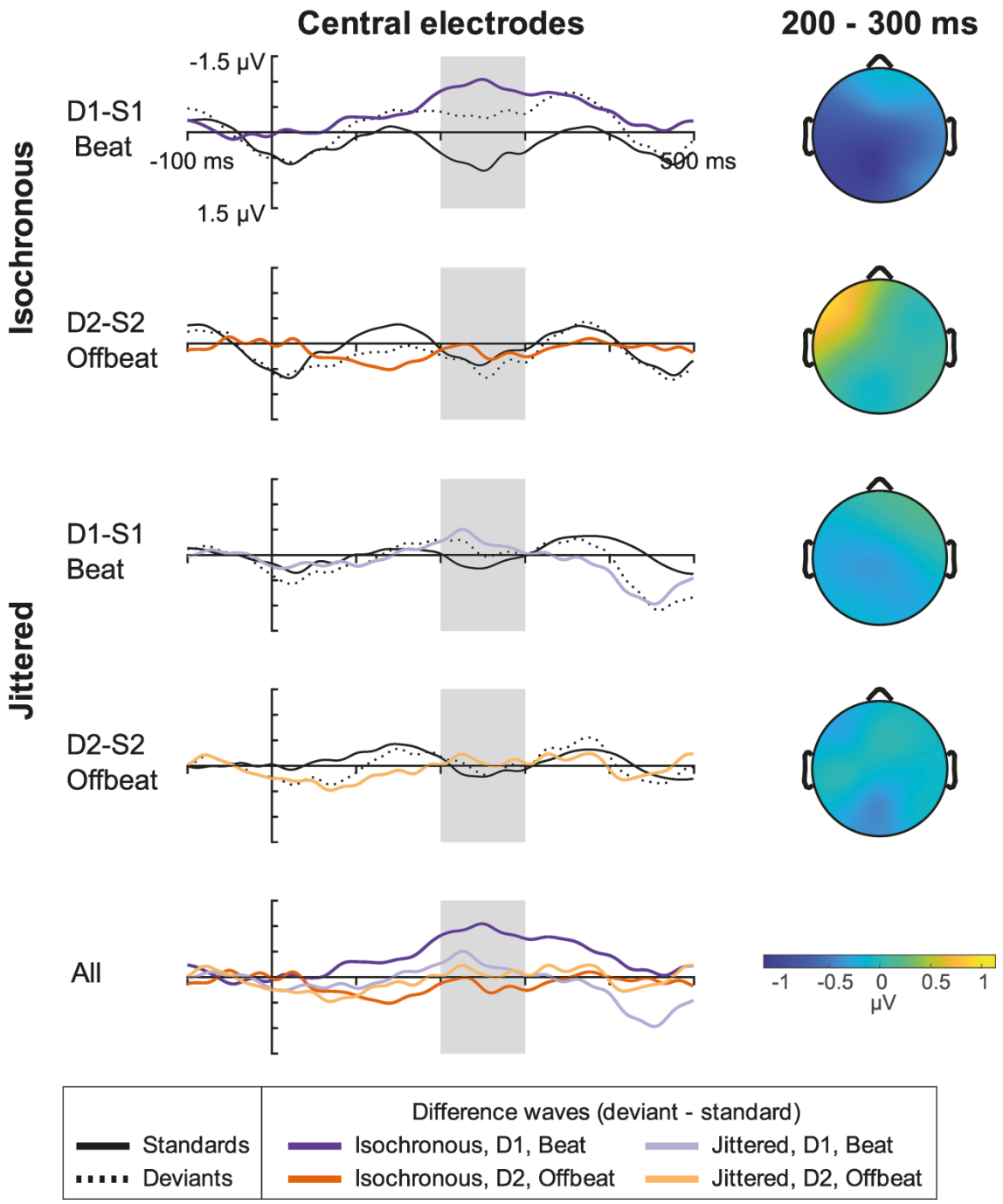
Figure 2



$$P = \begin{array}{c} \\ A \\ U \\ T \end{array} \begin{array}{c} \begin{array}{ccc} A & U & T \end{array} \\ \begin{bmatrix} .3 & .6 & .1 \\ .9 & 0 & .1 \\ .5 & .5 & 0 \end{bmatrix} \end{array}$$

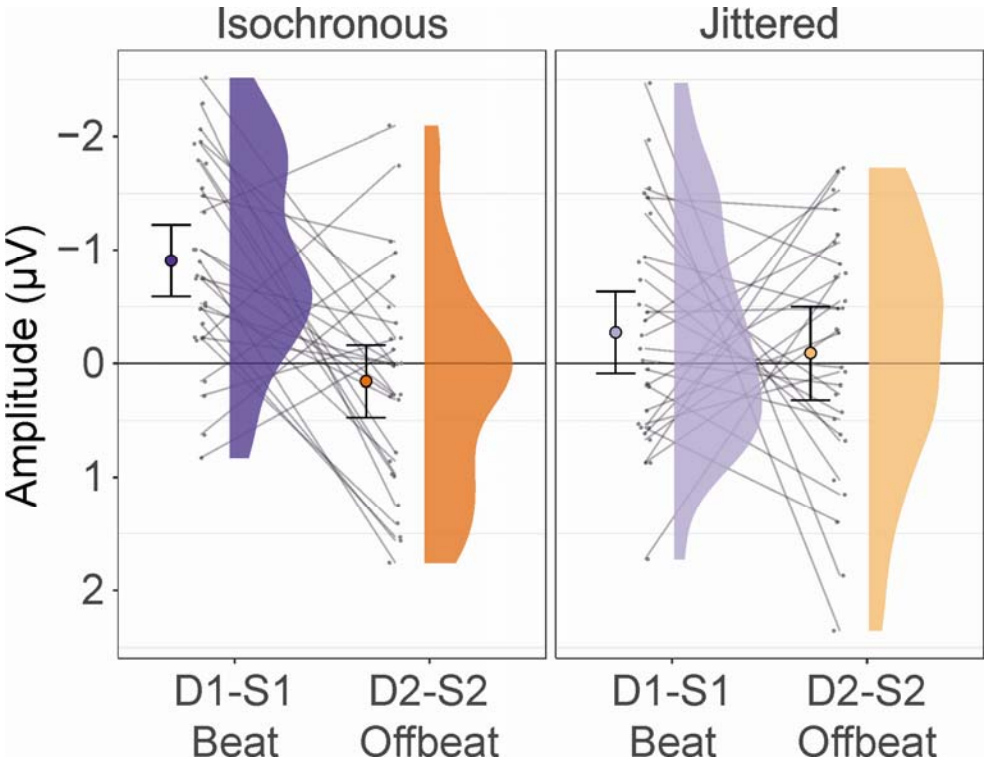
bioRxiv preprint doi: https://doi.org/10.1101/2022.12.20.521245; this version posted December 21, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

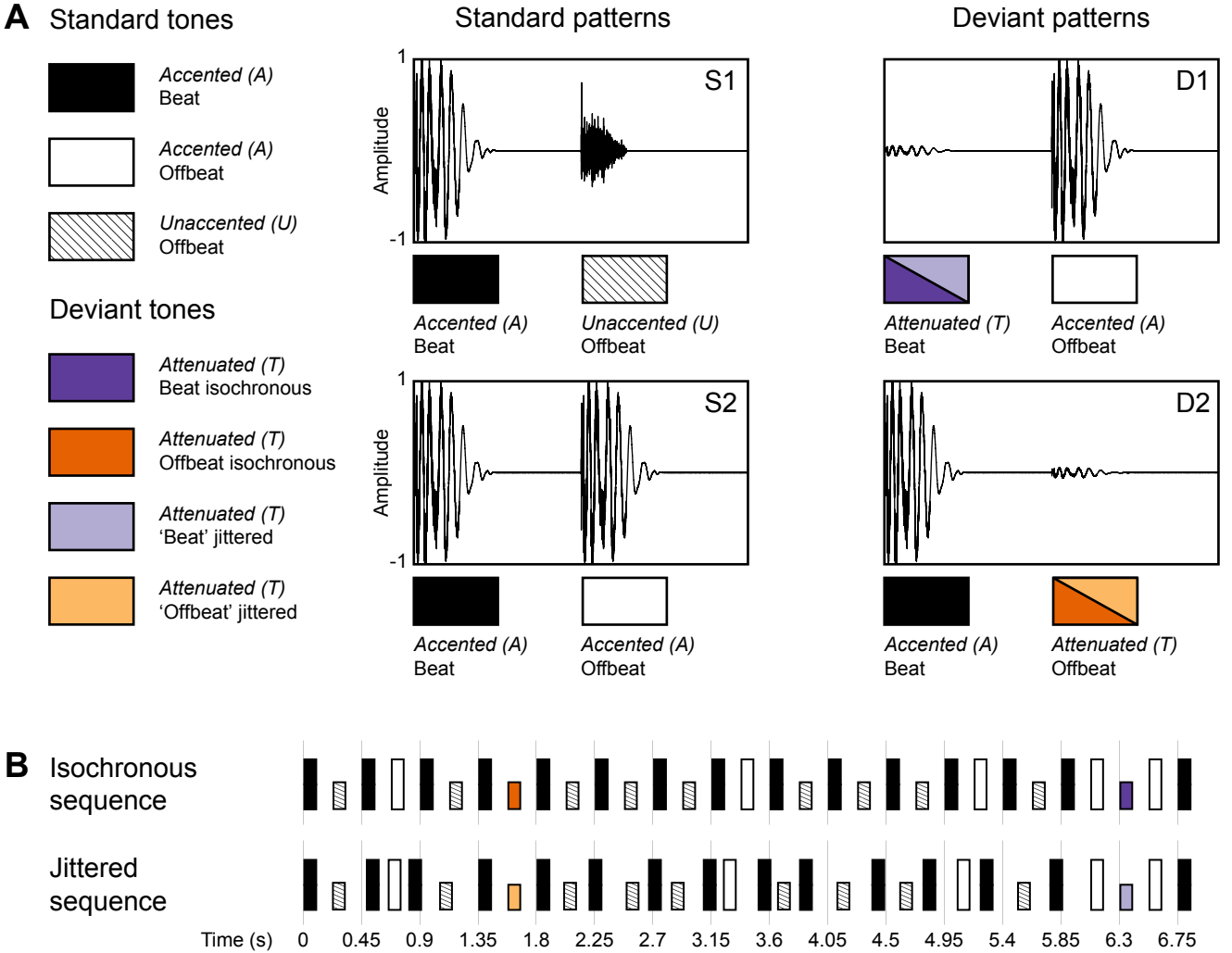
Figure 3

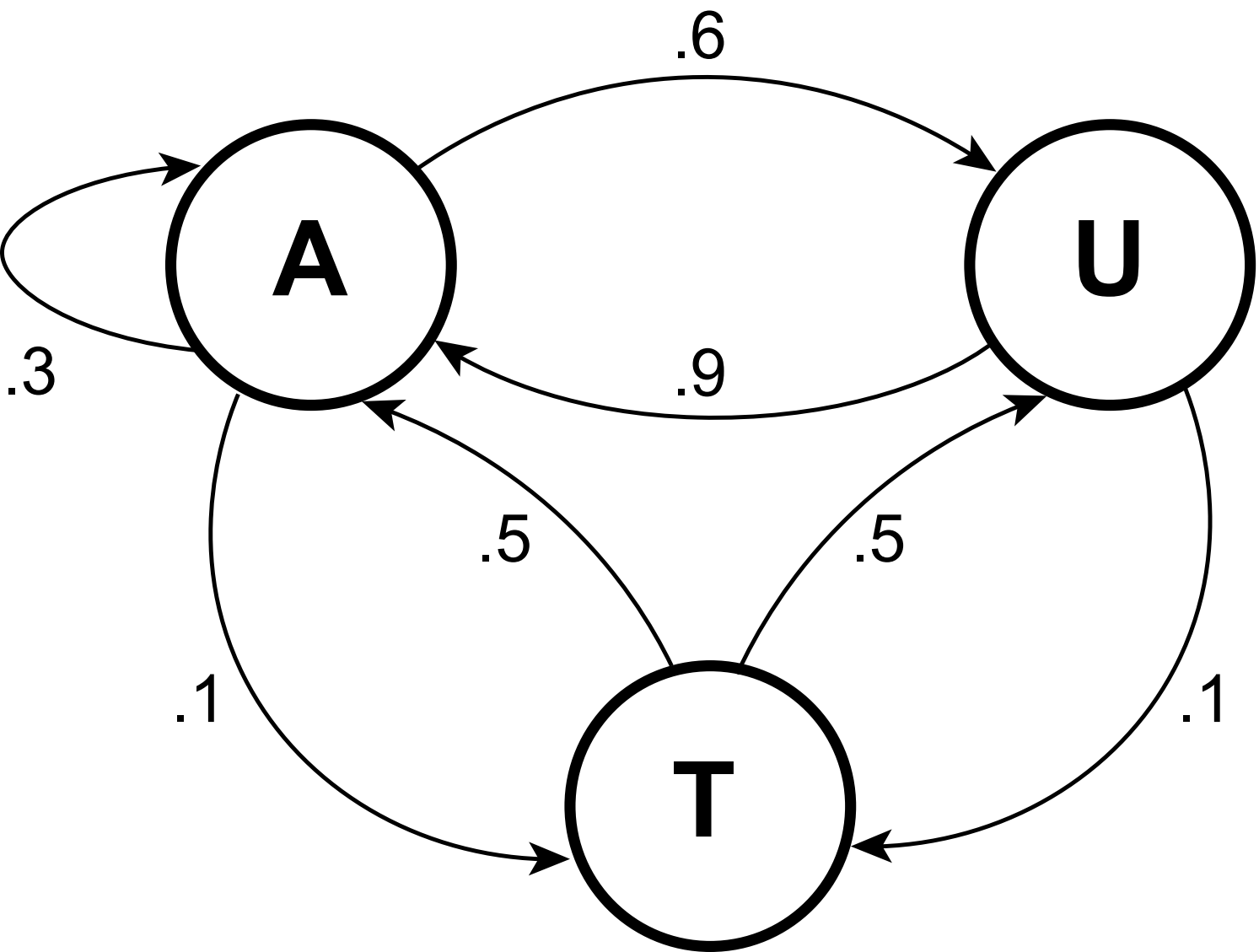


bioRxiv preprint doi: https://doi.org/10.1101/2022.12.20.521245; this version posted December 21, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Figure 4

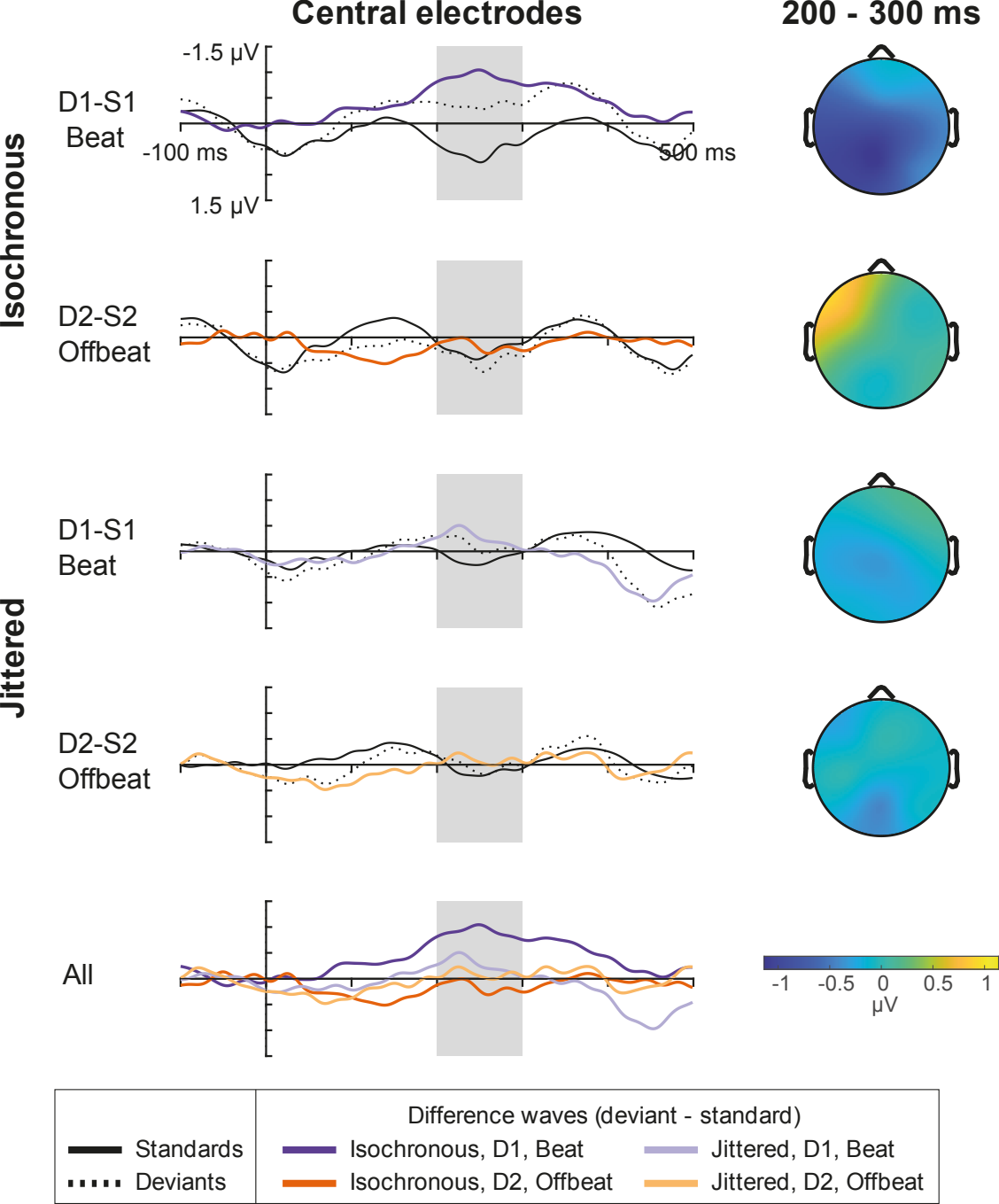


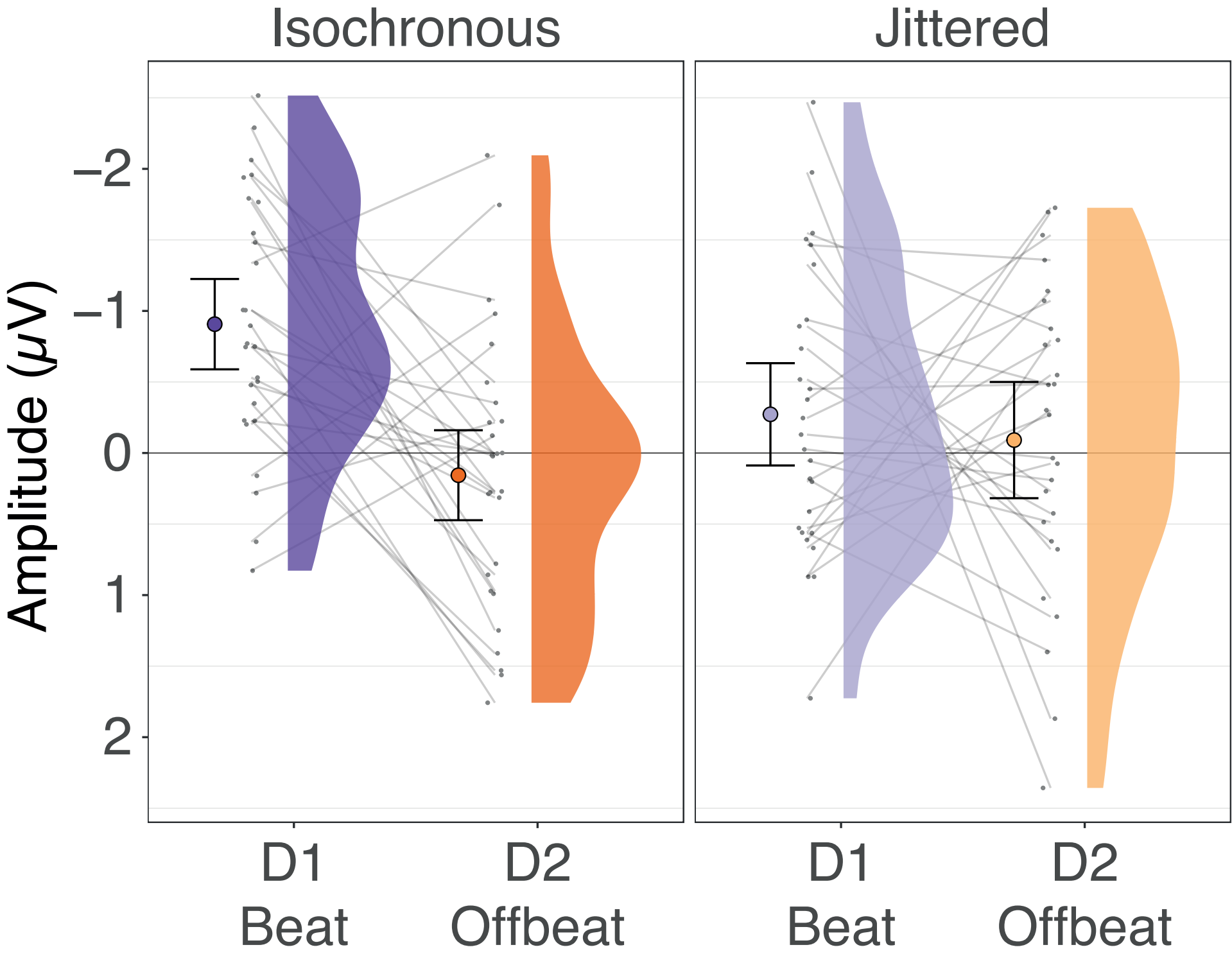

A
Standard tones
Accented (A)
Beat
Accented (A)
Offbeat
Unaccented (U)
Offbeat
Deviant tones
Attenuated (T)
Beat isochronous
Attenuated (T)
Offbeat isochronous
Attenuated (T)
'Beat' jittered
Attenuated (T)
'Offbeat' jittered
Standard patterns
Deviant patterns
S1
D1
S2
D2
Amplitude
1
-1
Accented (A)
Beat
Unaccented (U)
Offbeat
Attenuated (T)
Beat
Accented (A)
Offbeat
Accented (A)
Beat
Accented (A)
Offbeat
Accented (A)
Beat
Attenuated (T)
Offbeat
B
Isochronous sequence
Jittered sequence
Time (s)
0
0.45
0.9
1.35
1.8
2.25
2.7
3.15
3.6
4.05
4.5
4.95
5.4
5.85
6.3
6.75



A = Accented
U = Unaccented
T = Attenuated

$$P = \begin{array}{c} \\ A \\ U \\ T \end{array} \begin{array}{c} \begin{array}{ccc} A & U & T \end{array} \\ \begin{bmatrix} .3 & .6 & .1 \\ .9 & 0 & .1 \\ .5 & .5 & 0 \end{bmatrix} \end{array}$$


Central electrodes
200 - 300 ms
Isochronous
Jittered
D1-S1
Beat
D2-S2
Offbeat
D1-S1
Beat
D2-S2
Offbeat
All
-1.5 µV
1.5 µV
-100 ms
500 ms
-1
-0.5
0
0.5
1
µV
Standards
Deviants
Difference waves (deviant - standard)
Isochronous, D1, Beat
Isochronous, D2, Offbeat
Jittered, D1, Beat
Jittered, D2, Offbeat


Isochronous
Jittered
Amplitude (μV)
−2
−1
0
1
2
D1
Beat
D2
Offbeat
D1
Beat
D2
Offbeat