NeuroImage 273 (2023) 120090

Contents lists available at ScienceDirect

NeuroImage

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

Individual neurophysiological signatures of spontaneous rhythm processing

A. Criscuolo [a], M. Schwartze [a], M.J. Henry [b,c], C. Obermeier [d,e], S.A. Kotz [a,e,*]

[a] Department of Neuropsychology & Psychopharmacology, Faculty of Psychology and Neuroscience, Maastricht University, Maastricht 6200 MD, the Netherlands
[b] Max Planck Institute for Empirical Aesthetics, Frankfurt am Main, Germany
[c] Department of Psychology, Toronto Metropolitan University, Canada
[d] BG Klinikum Bergmannstrost Halle, Halle 06112, Germany
[e] Department of Neuropsychology, Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig 04103, Germany

A B S T R A C T

When sensory input conveys rhythmic regularity, we can form predictions about the timing of upcoming events. Although rhythm processing capacities differ considerably between individuals, these differences are often obscured by participant- and trial-level data averaging procedures in M/EEG research. Here, we systematically assessed neurophysiological variability displayed by individuals listening to isochronous (1.54 Hz) equitone sequences interspersed with unexpected (amplitude-attenuated) deviant tones. Our approach aimed at revealing time-varying adaptive neural mechanisms for sampling the acoustic environment at multiple timescales. Rhythm tracking analyses confirmed that individuals encode temporal regularities and form temporal expectations, as indicated in delta-band (1.54 Hz) power and its anticipatory phase alignment to expected tone onsets. Zooming into tone- and participant-level data, we further characterized intra- and inter-individual variabilities in phase-alignment across auditory sequences. Further, individual modeling of beta-band tone-locked responses showed that a subset of auditory sequences was sampled rhythmically by superimposing binary (strong-weak; S-w), ternary (S-w-w) and mixed accentuation patterns. In these sequences, neural responses to standard and deviant tones were modulated by a binary accentuation pattern, thus pointing towards a mechanism of dynamic attending. Altogether, the current results point toward complementary roles of delta- and beta-band activity in rhythm processing and further highlight diverse and adaptive mechanisms to track and sample the acoustic environment at multiple timescales, even in the absence of task-specific instructions.

1. Introduction

Due to the inherently rhythmic nature of many environmental stimuli, neurocognitive functions such as attention (Lakatos et al., 2008) sensorimotor behavior (Merker et al., 2009), speech (Giraud and Poeppel, 2012; Kotz and Schwartze, 2010), reading (Goswami, 2011), and music processing (Doelling and Poeppel, 2015) rely on basic timing capacities. To generate a temporally coherent representation of a rhythmic environment, we track stimulus periodicities, use smart grouping, and continuously segment and combine multiple inputs in time (Buzsáki, 2009; Schroeder and Lakatos, 2009; Thut et al., 2012a; Zoefel and VanRullen, 2016). According to the dynamic attending theory (Large and Jones, 1999) these processes can reflect how internal attending rhythms synchronize with external rhythms. This and similar theoretical views (Fraisse, 1963; p. 18) suggest that oscillatory brain activity instantiates a realistic model for "adaptation by anticipation". Accordingly, temporally regular sensory input should make future events predictable and improve the overall effectiveness of behavior. Thus, allocating attention to salient events can facilitate sensory processing, perception, and action (Friston, 2005; Arnal, 2012; Schröger et al., 2015; Koelsch et al., 2019).

However, *continuous change is a fundamental characteristic of life* (Schwartze et al., 2011), and next to temporal regularities, we frequently encounter irregular rhythms or sudden environmental changes. To account for these dynamics, any realistic adaptation mechanism likely tolerates a certain degree of temporal irregularity or unpredictability while trying to achieve synchronization (Barnes and Jones, 2000). Endogenous oscillatory activity must hence not only be precise and stable over time, but also flexible enough to achieve adequate adaptive timing (e.g., speeding-up or slowing-down). Oscillatory brain activity can actively track and process (quasi-)periodic and never strictly isochronous signals such as speech that includes rhythmic variations at phoneme (25–35 Hz), syllable (4–8 Hz), or word (1–3 Hz) rates, as well as slower fluctuations (>1 Hz) reflecting linguistic boundaries (Ding et al., 2015; Giraud and Poeppel, 2012). Moreover, it can rapidly adapt to changes in the sensory environment, likely through phase resetting (Barnes and Jones, 2000; Haegens and Zion Golumbic, 2018; Mormann et al., 2005; Obleser et al., 2012; Zoefel et al., 2018).

Particularly, delta- (δ; 1–4 Hz) and beta- (β; 12–25 Hz) frequency oscillations have been associated with rhythm processing, temporal prediction, and attention in humans (Arnal, 2012; Biau and Kotz, 2018; Colling et al., 2017; Fujioka et al., 2012, 2015; Morillon et al., 2016; Nozaradan et al., 2015, 2017a) and in non-human primates (Bartolo and Merchant, 2015; Merchant et al., 2015; Merchant and Bartolo, 2018; Patel and Iversen, 2014). However, prior behavioral studies on temporal processing also reported high within-subject variability (Baath, 2015;

* Corresponding author at: Department of Neuropsychology & Psychopharmacology, Faculty of Psychology and Neuroscience, Maastricht University, Maastricht 6200 MD, the Netherlands.
E-mail address: sonja.kotz@maastrichtuniversity.nl (S.A. Kotz).

https://doi.org/10.1016/j.neuroimage.2023.120090.
Received 22 September 2022; Received in revised form 14 March 2023; Accepted 4 April 2023
Available online 5 April 2023.

Poudrier, 2020). Thus, a critical question arises: can the systematic assessment of individual neurophysiological variability improve our understanding of how we process environmental rhythms? (Grahn and McAuley, 2009; Kononowicz and van Rijn, 2015; Nave et al., 2022; Waschke et al., 2021).

To address this question, we let participants listen to isochronous (1.54 Hz) equitone sequences, comprising frequent standard and either one or two amplitude-attenuated deviant tones while EEG was recorded. The small amplitude attenuation represented a minimally distracting unpredictable deviation from the established regularity of the auditory sequence. A deviant-counting task focused participant's attention on an acoustic feature (loudness) of the stimuli, diverting attention away from the temporal properties of the auditory sequence. We first assessed the neural signatures of spontaneous temporal processing by means of rhythm tracking analyses, expecting to observe a peak of neural oscillatory power at the stimulation rate (1.54 Hz, i.e., delta-band), and a consistent phase-alignment towards expected tone onsets, indicating temporal prediction.

Second, we focused on the known human disposition to superimpose accentuation patterns onto isochronous equitone sequences as exemplified by the "tick-tock" clock effect (Brochard et al., 2003). This typically results in perceived binary (strong-weak (S-w)) accentuations, while ternary (S-w-w) and other accentuations are possible (Abecasis et al., 2005; Baath, 2015; Brochard et al., 2003; Fujioka et al., 2015; Polak et al., 2018; Poudrier, 2020; Savage et al., 2015).

Informed by previous results that confirmed a role of beta oscillations in temporal predictions (Fujioka et al., 2012) and beat processing (Fujioka et al., 2015), we expected that such accentuation patterns would show in the envelope of beta-band activity. We employed a fixed stimulation frequency where we expected more binary than ternary accents (Abecasis et al., 2005; Baath, 2015; Brochard et al., 2003; Fujioka et al., 2015; Poudrier, 2020).

Third, we looked into individual differences to answer whether: (i) individuals always accentuate, (ii) individuals accentuate consistently over time, and (iii) accentuation patterns modulate cognitive processes as reflected in deviance processing (Brochard et al., 2003). To this end, we modelled single-participant and single-trial time-locked beta-band fluctuations to gain a better understanding of intra- and inter-individual neurophysiological variability indicating individual mechanisms employed to sample, evaluate, and adapt to environmental rhythms.

2. Materials & methods

2.1. Participants

Twenty native German speakers participated in the study and signed written informed consent in accordance with the guidelines of the ethics committee of the University of Leipzig and the declaration of Helsinki. Participants (9 females; 21–29 years of age, mean 26.2 years) were right-handed (mean laterality coefficient 93.8, Oldfield, 1971), had normal or corrected-to-normal vision, and no known hearing deficits. Participants received 8€/h for taking part in the study. Participants were not asked to indicate musical expertise and/or daily music listening.

2.2. Experimental design and procedure

The stimuli comprised 192 sequences, consisting of 13-to-16 tones (F0 = 400 Hz, duration = 50 ms, amplitude = 70 dB SPL; standard STD), presented in two recording sessions. One or two deviant tones (DEV), attenuated by 4 dB relative to the STD tones, were embedded in each sequence, replacing STD tones. The first DEV tone could either occur in an odd or even-numbered position (8–11th), corresponding to a hypothetical binary Strong-weak (S-w) accentuated position, while the second DEV always fell on the 12th position (w position, Fig. 1). The inter-onset-interval between successive tones was 650 ms, resulting in a fixed stimulation frequency of 1.54 Hz and a total sequence duration of 8.45–10.4 s (13 to 16 tones * 650 ms). This paradigm was thus comparable to previous behavioral studies on subjective accentuation (Brochard et al., 2003; Poudrier, 2020).

Participants were seated in a dimly lit soundproof chamber facing a computer screen. Every trial started with a fixation cross (500 ms), followed by the presentation of the tone sequence. The cross was continuously displayed on the screen, preventing excessive eye movements during the presentation of the tone sequences. At the end of each sequence, a response screen appeared and prompted participants to immediately press a response button to indicate whether they had heard one or two softer tones. The button assignment was counterbalanced across participants. After the response, there was an inter-trial interval of 2000 ms. A session was divided into two blocks of approximately 10 min each, with a short pause in between (about 25 min total duration).

2.3. EEG recording

The EEG was recorded from 59 Ag/AgCl scalp electrodes (Electrocap International), amplified using a PORTI-32/MREFA amplifier (DC to 135 Hz), and digitized at 500 Hz. Electrode impedances were kept below 5 kΩ. The left mastoid served as online reference. Additional vertical and horizontal electro-oculograms (EOGs) were recorded.

2.4. Data analysis

2.4.1. Behavioral analysis

Behavioral data (i.e., response accuracy) were analyzed with a repeated-measures ANOVA with the deviant position (odd vs. even) as the independent variable and sequence order (position in the sequence, e.g., 8th or 9th) as a covariate.

2.4.2. EEG preprocessing

Data were pre-processed using combined custom Matlab scripts/functions and the Matlab-based FieldTrip toolbox (Oostenveld et al., 2011). Data were first re-referenced to the average of the two mastoid electrodes and then band-pass filtered with a 4th order Butterworth filter in the frequency range of 0.1–50 Hz (*ft_preprocessing*). Eye-blinks and other artifacts were identified using independent component analysis ('*fastICA'* implemented in FieldTrip). A semi-automated pipeline was used to identify EEG components with a strong correlation (>0.4; labeled as "bad" components) with the EOG time-courses to inspect the respective topographical distribution across scalp electrodes and to remove "bad" components. Data segmentation was then conducted separately for the rhythm-tracking, event-related potential (ERP), and time-frequency representation (TFR) analyses. Note that behavioral and rhythm tracking analyses are independent of the modeling and the analyses on individual accentuation patterns described later.

2.4.3. Rhythm tracking analyses

Rhythm-tracking analyses involved neural responses to the full equitone sequences. Following ICA, 192 (96 sequences * 2 sessions per participant) segments were created, starting from the first tone onset up to the 13th tone offset (8.45 s). Fast-Fourier transform and rhythm tracking analyses, however, were performed on shorter segments starting from the 3rd up to the 13th tone offset (7.15s-long). The first two tones of the sequence were excluded from further analyses as it is known that they elicit much stronger event-related responses than tones in later positions of the sequence.

Next, we selected a fronto-central channel cluster, encompassing the sensor-level correspondents of prefrontal, pre-, para-, and post-central regions that were highlighted in previous MEG studies which employed source-localization analyses (Fujioka et al., 2012, 2015). The cluster included 16 channels: 'AFz', 'AF3′, 'AF4′, 'F3′, 'F4′, 'F5′, 'F6′, 'FCz', 'FC3′,

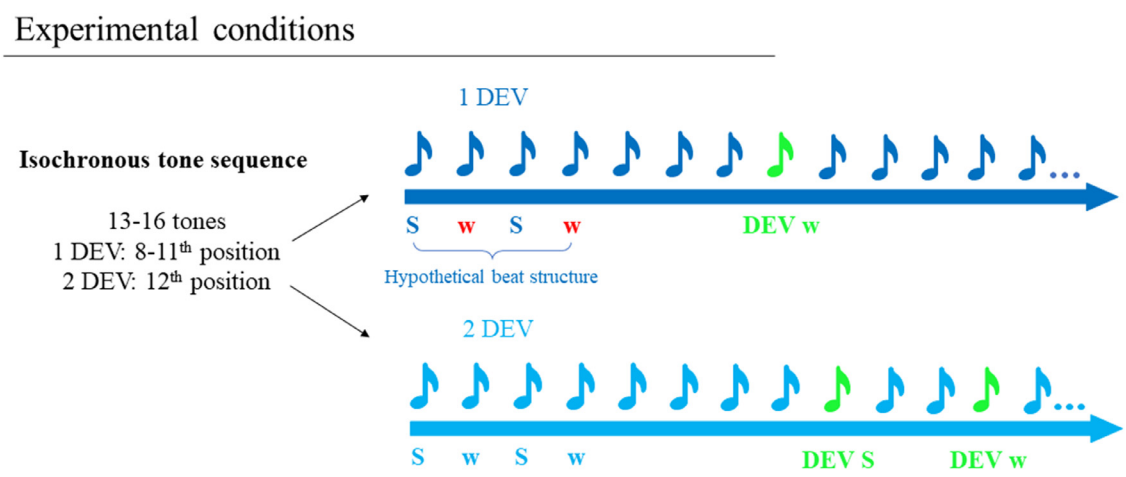

Fig. 1. Experimental conditions.
Participants listened to 192 isochronous tone sequences, containing 13-to-16 tones and either one or two deviants (DEV). The first DEV could either fall on positions 8,9,10,11th, while the second DEV always fell on position 12th. A hypothetical binary accentuation pattern would designate adjacent tones as Strong-weak (S-w) duplets. In this case, the first DEV would occur with equal probability in S-w positions.

'FC4′, 'FC5′, 'FC6′, 'C1′, 'C2′, 'C3′, 'C4′. Data from this fronto-central cluster were not averaged at this stage and were used for Fast-Fourier transform (FFT) and phase-locking analyses. While the FFT analysis mirrors the well-known 'frequency-tagging' approach (Nozaradan et al., 2017), the employed phase analyses provide a richer description of how neural oscillations *track* external auditory rhythms. Hence, the name 'rhythm tracking analyses'.

Fast-Fourier transform. Single-trial data from the fronto-central cluster were submitted to a FFT ("FFT data") with an output frequency resolution of 0.14 Hz (1/7.15 s = 0.14 Hz). Spectral power was calculated as the squared absolute value of the complex Fourier output. Each trial was then normalized by the standard deviation across trials, per frequency-bin and channel. Lastly, the frequency-domain data were averaged across channels and trials. For illustration purposes only, the Fourier spectrum was restricted to 1–4 Hz (Fig. 2A). The complex Fourier spectrum was also used to calculate inter-trial phase coherence (ITPC; Fig. 2A). This was calculated by dividing the Fourier coefficients by their absolute values (thus, normalizing the values to be on the unit circle), calculating the mean of these values, and finally taking the absolute value of the complex mean. Further documentation can be found on the FieldTrip website (https://www.fieldtriptoolbox.org/faq/itc/).

Phase-locking analyses. A time-resolved phase-locking analysis was performed to estimate the phase relationship between neural activity at the stimulation frequency and the sequential tone onsets.

The 8.45s-long data segments from the fronto-central cluster were bandpass-filtered with a 4th order Butterworth filter around the stimulation frequency (1.04–2.04 Hz, obtaining a 1.54 Hz center frequency; *ft_preprocessing*) and Hilbert-transformed to extract the analytic signal. The time-course of the real part of the analytic signal was then plotted (Fig. 2B) as a function of the STD tone onset preceding (blue) and following (red) the DEV (green; note that this plot serves an illustrative purpose only). Phase-locking analyses focused on tones 3 to 13, and were performed at the sequence and channel levels by means of circular statistics (circular toolbox in Matlab; (Berens, 2009), based on the circular mean phase-angles estimated in the ~60 ms (i.e., a time-window proportional to the stimulation frequency = 1/1.54 Hz/10) preceding individual tone onsets. Next, the sequence- and channel-level mean vector length were calculated (MVL; (Berens, 2009) for pre-DEV STD tones and the resultant values then pooled across channels. The focus on pre-DEV tones only is motivated by the fact that the onset of a DEV tone might disrupt the predictability of the auditory sequence, and further induces a phase-reset of oscillatory activity. Next, MVLs for pre-DEV STD tones were statistically assessed against the MVL from a random distribution (random uniform distribution of phase-angles) by means of 1000 permutation tests. A p-value lower than 0.05 was considered statistically significant. For illustrative purposes, we also calculated participant-, channel- and sequence-level 'relative phase angles': these were expressed as the absolute phase difference between phase-angles for each channel and tone position (e.g., 3rd to 8th) and the most common phase-angle in the sequence. The most common phase-angle was identified by means of the 'histogram' function in MATLAB, using 'probability' as parameter after rounding phase values to 1 decimal. This means, a probability value is attached to each of the phase-angles within a single-participant, -channel, and -sequence, and across a tone positions (3rd to 8th). Next, the phase value with the highest probability (i.e., the most common) was used as a reference to calculate the 'relative phase-angles'. Thus, we computed the absolute phase difference between each of the phase-angles and the most common phase value.

Examples of participant- and sequence-level relative phase-angles are plotted in Fig. 2C, and the pooling over participants, sequences, and channels is provided in Fig. 2D.

2.4.4. ERP and TFR data

After ICA, data were segmented into 4s-long epochs symmetrically time-locked to every tone onset. Next, we employed an automatic channel-by-channel, trial- and participant-level artifact suppression procedure (comparable to Kaneshiro et al., 2020). Artifact suppression focused on time-windows ranging from −.4 to 0.4 s relative to each stimulus onset. Amplitude values were temporarily normalized by their standard deviation across trials and outliers (data points per epoch and channel) were defined by means of a threshold criterion (values > mean + 4*SD). The identified noisy time-windows (with 50 ms symmetrical padding) were then filled with NaNs, and these missing values were replaced by means of cubic temporal interpolation (*'pchip'* option for both the built-in Matlab and FieldTrip-based interpolation functions) considering the time-course of neighboring time-windows (extending up to 100 ms when possible, automatically reduced otherwise). The current approach is a novel data-driven procedure developed to minimize the data loss. It differed from (Kaneshiro et al., 2020) insofar as the channel-by-channel routine allowed the algorithm to flexibly adapt the outlier threshold estimates to the inherent noise varying over channels. Descriptive analyses revealed that the artifact suppression procedure was used for 5% of trials on average, on time-windows 100 ms long, and most likely between 350–400 ms after stimulus. Critically, this strategy

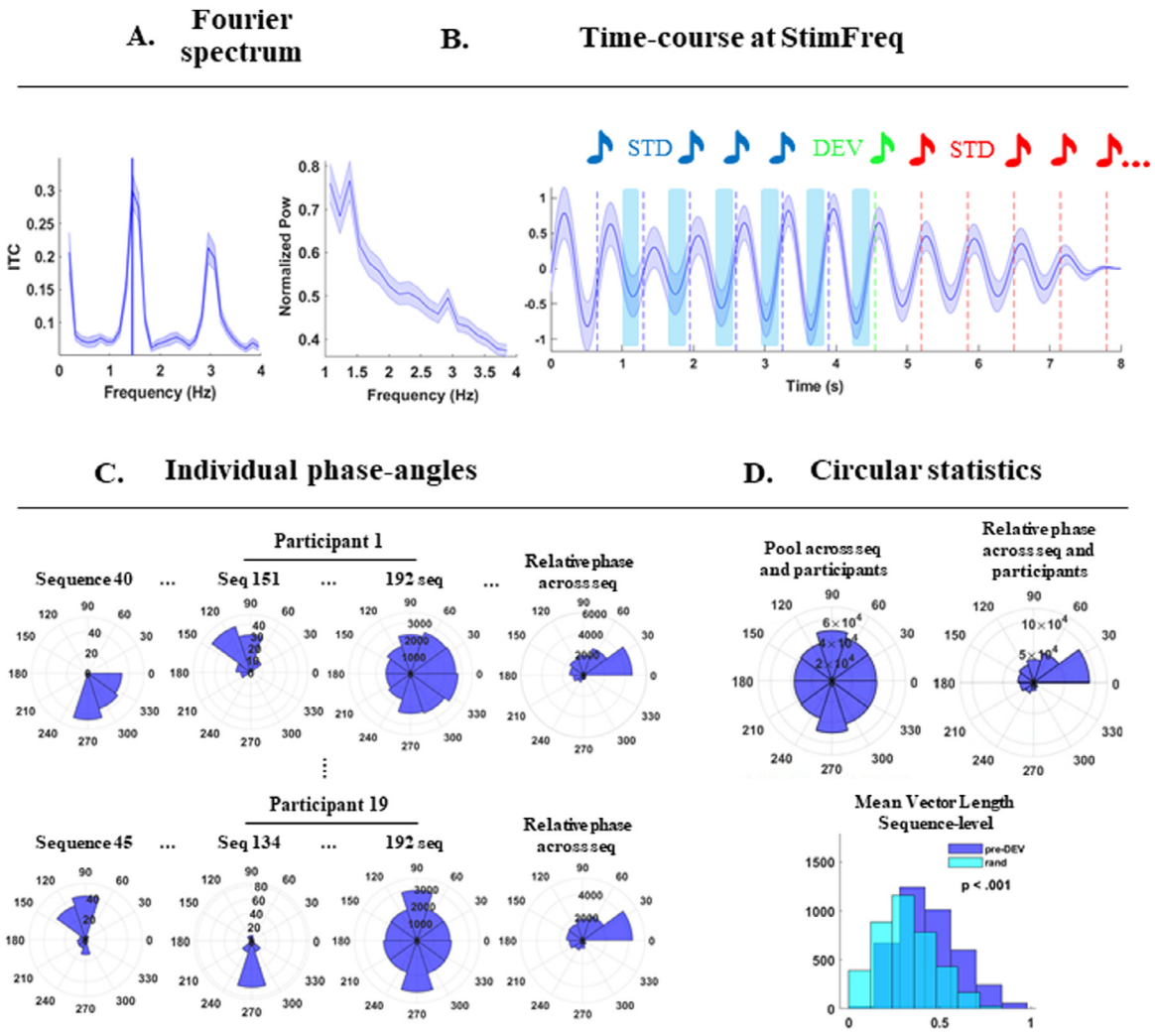

Fig. 2. Rhythm tracking analyses.

A: Fourier spectrum of neural activity along the entire equitone sequence. The plots display, in order, the inter-trial phase coherence (ITC; left) and the grand-average normalized power (right) in the frequency range from 1 to 4 Hz. The vertical line highlights the peak of phase coherence at the stimulation frequency (1.54 Hz). B: time-course of neural activity at the stimulation frequency. Vertical bars indicate the onsets of STD tones prior- (blue) and post-DEV (red). The DEV onset is reported in green. Blue shades represent the standard errors. Light-blue rectangles indicate the focus on the pre-stimulus intervals (not scaled). C: Polar histograms for single-participant and sequence-level phase angles extracted from 60 ms prior to the onsets of STD tones prior (blue) to the DEV from the fronto-central cluster of interest. Here, we report a few sequence-level phase-angles from Participant 1 (top) and Participant 19 (bottom). On the right, the polar histograms report the distribution of phase-angles across all trials (192 per participant) and the 'relative phase' across sequences. This is a measure of deviation from the most common phase-angle, at the sequence-level. D. Group-level phase-angles are randomly distributed around the polar histogram. On its right, the group-level 'relative phase'. These phase-angles indicate a variation from the most common phase-angle. At the bottom, the distribution of mean vector length (MVL) calculated at the single-participant and sequence-level and averaged across the fronto-central cluster of interest. Importantly, these MVLs are based on the raw phase-angles for pre-DEV (blue) and are statistically compared to the MVL for random distribution of phase-angles. Single-participant statistics are reported in Suppl. Table 1.

allowed keeping all trials instead of rejecting entire epochs only partially contaminated by artifacts (i.e., as it usually happens in the typical standard artifact rejection procedure) (Kaneshiro et al., 2020).

Next, a standard whole-trial rejection procedure based on an amplitude criterion (85uV) was applied. Data were then segmented for event-related-potential (ERP) analyses ("ERP data"), including 500 ms prior and following each tone onset (1 s in total). Data for the time-frequency representation analyses ("TFR data") were not further segmented at this stage. ERP data were band-pass filtered between 1–30 Hz, while TFR data were low-pass filtered at 40 Hz. Data were downsampled to 250 Hz.

Single-trial TFR data underwent time-frequency transformation by means of a wavelet-transform (Cohen, 2014) with a frequency resolution of 0.25 Hz. The number of fitted cycles ranged from 3 for the low frequencies (<5 Hz) to 10 for high frequencies (>5 Hz and up to 40 Hz). The single-trial approach results in 'induced' (as compared to 'evoked') responses. TFR data were then re-segmented to reduce the total length to 2 s, symmetric around tone onsets.

Mean correction of ERP and TFR data. Single-trial ERP amplitudes were mean-corrected by a global average over all epochs and computed in a time-window ranging from −0.2 to 0.3 s relative to tone onset. Similarly, single-trial TFR power was normalized by computing relative percent change with reference to the global mean power across epochs (−0.2 to 0.3 s relative to tone onset). This previously applied approach (Abbasi and Gross, 2020; Fujioka et al., 2012) was preferred over classical baseline correction because we aimed at analyzing power fluctuations in pre-stimulus intervals. Finally, we calculated a fronto-central channel cluster average (using the same channels as for the rhythm tracking analyses). All subsequent analyses were performed exclusively on this channel cluster.

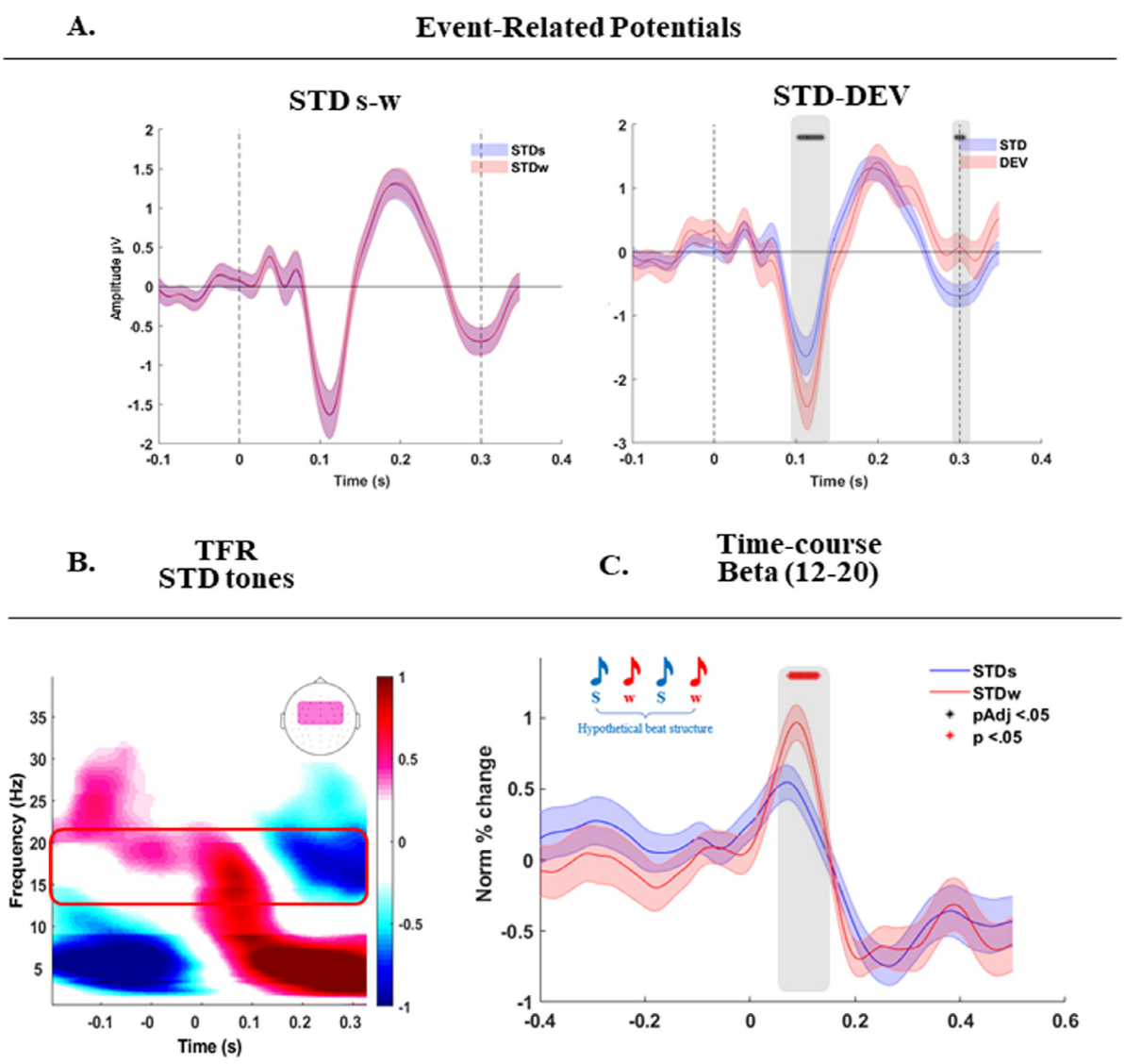

Fig. 3. Binary accents on hypothetical S-w positions.
A: On the left, ERP responses for STD tones in S (blue) w (red) positions. On the right, ERP responses for STD (blue) and DEV tones. Stars indicate significant time-windows, as assessed by means of paired-sample t-tests (FDR-adjusted p <0.05). B: grand-average time-frequency spectrum time-locked to STD tones (−.2 to 0.35 s). The frequency range spans from 1-to-40 Hz with a frequency resolution of 0.25 Hz. The red rectangle highlights evoked responses in the low-beta (12–20 Hz) frequency range, on which we performed statistical comparisons in C. The topographic plot on top displays the FC cluster average in use. C: extracted time-course of low-beta activity in hypothetical S-w positions, time-locked to STD tones onsets, in blue for odd-numbered positions (Strong) and red for even-numbered positions (weak). Shaded colors report standard errors. On top, a gray rectangle delineates the time-window in which statistical testing without multiple-comparison correction showed a difference between S and w positions. The comparison, however, did not survive FDR correction.

ERP analyses. Evoked responses over trials were averaged separately for STD and DEV tones and for odd (hypothetical "Strong" position in a binary accentuation pattern; S, Fig. 3A, left) and even ("weak"; w) positions from the 3rd to the 11th tone. The first STD tones were disregarded to exclude the increased responses typically observed at the beginning of an auditory sequence. Fig. 3A shows the respective ERPs for the averaged fronto-central cluster, for STD tones in hypothetical strong and weak positions (left) and for the comparison of STD and DEV tones averaged over these positions (right). Statistical analysis was performed by means of paired-sample t-tests over a time-window ranging from 0 to 350 ms relative to tone onset. An FDR-adjusted p-value lower than 0.05 was considered statistically significant (Benjamini & Hochberg correction).

TFR analyses. Time–frequency representations were averaged over STD trials, separately for odd and even positions (hypothetical strong and weak positions, respectively; Fig. 3B). Mean amplitudes in the low-beta band (low-β; 12–20 Hz; Biau and Kotz, 2018) were then statistically compared for S-w positions by means of paired-sample t-tests over a time-window ranging from 0 to 350 ms relative to tone onsets. An FDR-adjusted p-value lower than 0.05 was considered statistically significant (Benjamini & Hochberg correction).

Individual classification of accentuations. An individual modeling approach was adopted to identify binary and ternary accentuation patterns. Similar to earlier studies on predictive timing, (Fujioka et al., 2012, 2015), we focused on single-participant's low-β mean power peaks for STD tones in the first eight positions of the equitone sequence. Tone-level mean power peaks were calculated as follows: we first located the power peaks in time-windows centered at 100 ms post-stimulus (resulting peaks from analyses in Fig. 3C) and extending 60 ms (proportional to the center frequency of interest; for low-β:

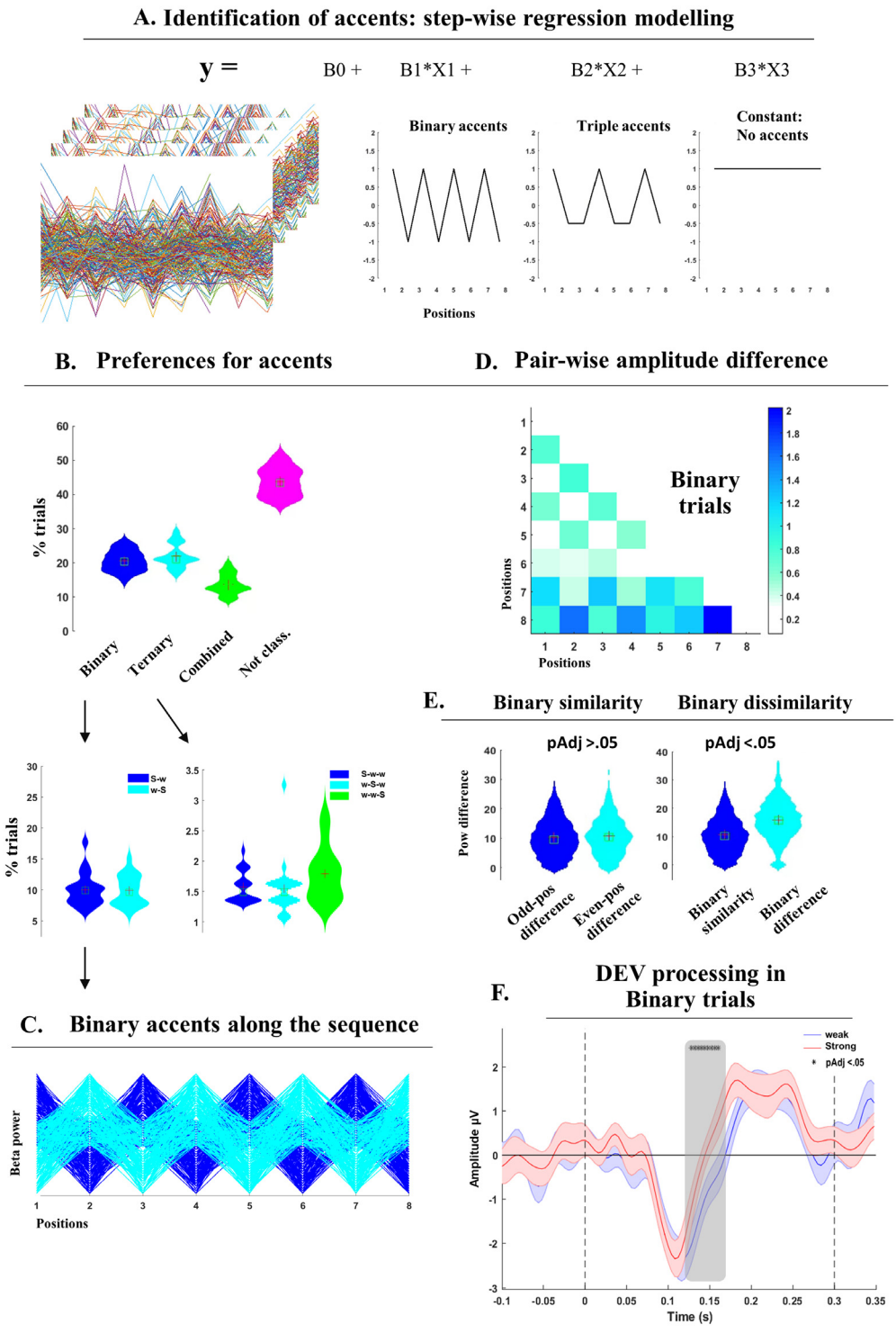

Fig. 4. Individual classification of accents and analyses on binary accents.
A: We modelled potential accentuation patterns by means of stepwise regression modeling and using low-beta post-stimulus responses as a dependent variable. The predictors were a binary (1, −1), a triple (1, −.5, −.5) and a constant term (ones). B: preferences for accents, as reported from the modeling. In order, we plot the distribution of trials assigned to binary, ternary, combined (binary-ternary) accents, and 'not classified' (neither binary nor ternary) across participants. At the bottom, we zoom into binary trials and distinguish S-w accents from w-S accents based on trial-level Beta coefficients from the modeling. Similarly, on its right side, the distribution of ternary trials showing S-w-w, w-S-w, and w-w-S accentuation patterns. To extract these three accentuation patterns, we performed separate stepwise regression modeling as explained in the method section. C: Exemplar S-w and w-S accent fluctuations expected along 8 positions of the auditory sequence in the 'binary' trials. Blue for S-w; cyan for w-S sequences. This plot has illustration purposes only. D: grand-average pair-wise difference for low-beta peaks across the first 8 positions of the auditory sequence in binary trials. E: on the left, the distribution of amplitude differences across odd-numbered positions (in blue) and even-numbered positions (cyan). The average of these two distributions forms the 'Binary similarity'. On the right, the 'binary similarity' (blue) and the mean amplitude difference of odd- versus even-numbered position ('binary dissimilarity'; in cyan). Statistical testing was performed by means of thousand permutation testing, and an FDR-adjusted $p < .05$ was considered as statistically significant. F: ERPs to DEV tones on S-w positions in the binary trials. Statistical testing reported a significant difference in the time-window between ~120–170 ms post-stimulus, as highlighted by the gray shades. ERPs to DEV tones in non-binary trials did not differ on S-w positions; see Suppl. Fig. 1.

1/16 Hz = 60 ms), prior and after each peak. Next, the resulting peak latencies were used to create a second time-window of interest centered around the individual peaks and extending 60 ms before and after it. Thus, tone-level mean power peaks were calculated within these 120ms-long time-windows following the stimulus onset (from now on, β-post).

Single-tone β-post were first concatenated to mimic an 8-tone sequence (i.e., a trial). Single-participant and trial-level β-post (eight tones) were then entered into a stepwise regression model (Fig. 4A) ('*stepwiselm*' in Matlab) with three predictors: a binary (values: 1, −1), a ternary (1, −.5, −.5), and a constant term (ones). The stepwise regression function searches for the predictor or a combination of predictors that maximizes the fit of the model to the real data. The model thus allows the combination of multiple predictors, but no interactions between terms. Once all combinations are tested, the winning model is chosen based on the estimated adjusted eta squared. Trials for which the winning model involved the binary predictor were labeled "binary", trials for which the winning model involved the ternary predictor were labeled "ternary". Accordingly, we interpreted (and labeled) the combination of binary and ternary terms as "combined". The remaining trials in which neither binary nor ternary predictors were included in the

winning model, were labeled as "not classified". Participant-level model results are provided in Suppl. Table 2 as "Preferences for accents" and expressed as the percentage of trials relative to the total number of auditory sequences (192 per participant). The subject-level goodness of fit of the model is provided in the same table. The "Preferences for accentuations" across participants are provided in Fig. 4B and expressed as the percentage of trials relative to the total number of auditory sequences (192 per participant).

Binary accents analyses. Further confirmatory analyses were performed to verify if the successfully identified "binary" trials indeed showed binary-like accentuation patterns. If so, neural responses to tones falling on odd-numbered positions should differ from those on even-numbered positions. However, there should be no differences for neural responses on the same positions: namely, tones falling on odd-numbered positions should elicit similar (i.e., non-significantly different) neural activity.

To verify if this was the case, we first vertically concatenated trials classified as "binary", and then computed a trial-based single-tone pairwise low-beta mean peak amplitude difference (corresponding lower-triangle 2-D means are provided in Fig. 4D), i.e., amplitude differences between responses to each tone in the sequence (1–8 positions). For example, the response amplitude for the 1st position was compared to the 2nd position, then to the 3rd, and so forth. In turn, the amplitude for the 2nd position was compared to the 3rd, the 4th etc. The resulting pairwise amplitude difference matrix had a size of N trials x N positions-1 x N positions-1. From this matrix, we statistically compared the pair-wise amplitude difference for tones in odd-positions (Fig. 4E; "odd-pos difference") and even-positions ("even-pos difference") by means of 1000 permutations of odd-even labels. An FDR-adjusted p-value lower than 0.05 was considered statistically significant (Benjamini & Hochberg correction). The two variables were then combined into a distribution of "binary similarity". The binary similarity combines the amplitude difference for tones in odd-numbered positions (1–3–5–7th) and the amplitude difference for tones on even-numbered positions (2–4–6–8th). Binary similarity was statistically compared to "binary dissimilarity", which was calculated as the mean difference of tones in odd versus even positions (Fig. 4E). Statistical testing was performed using 1000 permutations of odd-even labels, and an FDR-adjusted p-value lower than 0.05 was considered statistically significant (Benjamini & Hochberg correction).

DEV analyses as a function of binary accents. For each participant, we isolated binary trials identified with the described "individual classification of accents" procedure and explored the relation between accentuations and DEV processing to explore whether ERP responses differed according to their accents (S-w).

Statistical analyses compared ERPs to DEV tones in even- (8,10th) versus odd-numbered (9,11th) positions (Fig. 4F). Notably, we considered possible "pure binary" and "phase-shifted binary" accents (or inverse binary), where the former corresponds to the typical S-w pattern in odd and even positions, and the second to the reverse w-S pattern. Indeed, individuals may start accentuating at different times along the auditory sequences, and hypothetical S-w positions may likely fall on either odd or even positions. The selection was informed by the beta coefficients associated with the identified binary trials: a positive coefficient would indicate "pure binary" (S-w), while a negative coefficient would correspond to the "inverse binary" (w-S). The distribution of "pure" and "inverse" binary is provided at the bottom left of Fig. 4B, expressed as percent of trials relative to the total number of auditory sequences (192 per participant).

Next, we re-ordered accented S-w positions according to individual binary processing (odd-numbered positions falling on S accents in "pure binary" but on w in "inverse binary"), and statistically compared their associated ERPs time-courses (Fig. 4F). Statistical testing was performed by means of paired-sample t-tests, and an FDR-adjusted p-value lower than 0.05 was considered as statistically significant (Benjamini & Hochberg correction).

DEV analyses in ternary and non-classified trials. We also tested whether a similar S-w effect would be observed for DEV processing in 'non-classified' trials. ERP responses for DEV tones falling on accented S (odd-numbered positions) and w (even-numbered) positions were pooled and statistically compared by means of paired-sample t-tests. An FDR-adjusted p-value lower than 0.05 was considered statistically significant.

Next, we focused on potential "ternary" trials. In this case, the accent can either fall on the first (S-w-w), second (w-S-w), or third (w-w-S) position. To disentangle these three accentuation patterns from the distribution of "ternary" trials, we ran a second stepwise regression model. The model featured three predictors to include the possible accentuation types: 1,−.5,−.5 (pattern 1), −.5,1,−.5 (pattern 2) and −.5,−.5,1 (pattern 3). Again, the model did not allow interaction terms and the winning model was chosen based on adjusted eta squared. The output of the model is provided in Fig. 4B (bottom right), as the percentage distribution of three accentuation patterns across participants. Importantly, other accentuation patterns are possible as a S-w-w pattern could be represented by a stair-case amplitude change (e.g., 1, −.75, −.25) or a shuffled version (e.g., 1, −.25, −.75). Considering that binary accentuations are usually prevalent (Abecasis et al., 2005; Baath, 2015; Poudrier, 2020), possibly due to a cognitive bias structuring tonal sequences into groups of two (Polak et al., 2018; Savage et al., 2015), and that the employed stimulation rate may preferentially induce binary rather than ternary accentuations (Baath, 2015; Poudrier, 2020), even when tapping to polyrhythms (Møller et al., 2021), we did not build other models to test all possible ternary accentuations, or any other accentuation patterns. Of note is also that the employed model only used the first eight tones of the equitone sequence. This avoids the onset of DEV tones in later positions, which may disrupt ongoing accentuation, but inevitably leaves only up to two periods of a ternary accent (as compared to four repetitions of a binary accent). Consequently, even two small amplitude fluctuations with superimposed noise (inherent in EEG recordings) may drive the 'ternary' classification, but these trials may not necessarily reflect a true ternary accentuation. Accordingly, we accepted that a portion of 'ternary' trials might not be classified with the 1,−.5,−.5, −.5,1,−.5 and −.5,−.5,1 pattern.

3. Data and code availability

The analysis code and the data in use here will be stored in an open repository and can be provided upon reasonable request by the corresponding author.

4. Results

4.1. Behavioral data

We tested whether the counting of deviant tones (DEV) differed for deviants in odd or even positions in the equitone sequence. The respective ANOVA with deviant position (odd vs. even) and sequence order as a covariate did not reveal a significant effect of deviant position ($F(1,71) = 1.115$, $p = .295$, *eta-square* = 0.16) nor a significant effect of sequence order ($F(1,69) = 0.02$, $p = .97$, *eta-square* = 0). This indicates that DEV counting performance did not differ in the equitone sequence.

4.2. Rhythm tracking

Participants listened to equitone sequences presented at a stimulation rate of 1.54 Hz, and we tested *whether* and *how* their neural activity would show idiosyncratic signatures of rhythm tracking. When individuals listen to these sequences, their neural activity reflects the

timing of external events (Fig. 2A-B). Indeed, ITPC analyses and the normalized power of the Fourier spectrum both showed a clear peak at the stimulation frequency (1.54 Hz; Fig. 2A), and the time-course of delta-band neural activity showed a tendency to align to tone onsets (Fig. 2B; this plot is for illustration purposes only). To quantify the consistency of anticipatory phase alignment to the expected tone onsets, we tested the phase consistency of delta-band (1.54 Hz) neural activity in a time-window preceding tones onset. Phase-locking analyses focused on the ~60 ms (proportional to the stimulation frequency: 1/1.54 Hz/10) prior to tone onsets. Single-participant trial-level mean vector lengths (MVL) of STD tones preceding a DEV (pre-DEV) revealed a consistent phase-relationship with STD onsets: the MVL significantly differed from a random distribution (Fig. 2D, bottom; pre-DEV in blue; Suppl. Table 1 for statistical results). However, we observed intra- and inter-individual differences: participants' delta-band activity did not always synchronize to tone onsets with the same phase relationship (Fig. 2C). Rather, a broad range of possible phase-lags was observed across trials, both at the level of single-participants (Fig. 2C right) as well as when pooling values across participants (Fig. 2D top-left). The distribution of single-participant phase-angles across trials accordingly did not differ from a random distribution (Suppl. Table 1). Phase-angles were consistent within a trial (MVL statistics in Fig. 2D bottom and Suppl. Table 1) but differed across trials. To further explore this variability, we computed a measure of 'relative phase'. This was calculated, at the single-participant, channel- and sequence-level as the absolute difference from each phase-angle within one sequence and the most common phase. The distribution of relative phase-angles across trials and participants shows variance which mostly ranges between 0–30° (Fig. 2D, right), supporting the MVL calculation. Thus, individuals show a predictive and consistent phase-alignment of delta-band activity to expected tone onsets. However, the specific phase for this alignment is variable across trials.

4.3. Analyses on accentuations

We tested whether participants' neural activity would sample the acoustic environment by superimposing binary accentuation patterns (S-w accents in odd-numbered versus even-numbered positions) onto the equitone sequences. We analyzed event-related responses (ERP) to STD tones in S and w positions, and further inspected the time-frequency representation of time-locked responses.

ERPs to STD tones in S-w positions did not statistically differ (Fig. 3A). However, DEV tones elicited stronger N100 responses compared to STD tones (FDR-adjusted $p < .05$; Fig. 3A, right), confirming the processing of an unpredicted deviant tone.

The time-frequency representation plots of neural activity in response to STD tones mainly showed two event-locked responses (Fig. 3B): one in the theta (4–8 Hz) and one in the low-beta (low-β; 12–20 Hz) frequency-band. Following our hypotheses and informed by previous work (Fujioka et al., 2012, 2015), we focused on the time-course of activity in the low-β range and compared event-locked power fluctuations for STD tones on odd (S) to even (w) numbered positions along the sequence (Fig. 3C), corresponding to hypothetical S-w binary accents (blue and red, respectively). We statistically compared the time-courses of low-β activity (Fig. 3C) for STD tones in hypothetical S-w positions. The comparison did not survive FDR correction for multiple comparisons. Both FDR-adjusted p-values and non-adjusted p are provided in Fig. 3C (in black and red, respectively).

In summary, neither ERP nor TFR analyses revealed a binary accent, as STD tones elicited similar responses when they occurred in odd- and even-numbered positions in the auditory sequence. Similarly, ERPs to DEV in odd- and even-numbered positions did not statistically differ (Suppl. Fig. 1, bottom). These observations seemingly contradict original findings (Brochard et al., 2003), but might result from a different experimental setup and processing pipeline. For instance, the choice of region and time-windows of interest for statistical analyses differ from the original study.

To better characterize the phenomenon of *subjective accentuation*, we decided to zoom into inter- and intra-individual differences in *when* and *how* accentuation patterns are superimposed onto the auditory sequences. Indeed, not only may individuals start to accentuate at different points along the sequence (i.e., not necessarily at the beginning), they may also do so differently over time (e.g., binary or ternary accents), or even not accentuate at all (Brochard et al., 2003). These alternatives were tested with a novel modeling approach.

4.4. Modeling of individual accentuations

To address the questions of (i) whether everyone accentuates in a consistent way, (ii) whether everyone always accentuates in the first place, and (iii) whether accentuation patterns influence DEV processing (Brochard et al., 2003), we focused on trial-level data and modelled various accentuation patterns. We used a trial-based stepwise regression model to classify participant-level single-trial beta-band neural responses as best reflecting clear binary and ternary accents, or the absence of a corresponding accentuation. The choice to focus on beta-band activity was informed by previous evidence (Fujioka et al., 2012, 2015). The model predicted tone-by-tone low-β-post peak power from three predictors: binary, ternary, and constant (no accents) terms (Fig. 3A). Resulting 'preferences for accentuations' and goodness of fit are reported in Suppl. Table 2 and summarized in Fig. 4B. Note that most trials (~60%) did not clearly reflect either binary or ternary accents. In the absence of perceptual reports, it remains open whether participants did not perceive accentuations, or whether the current analyses are not sensitive enough to detect imagined accentuations at the level of neural activity.

Next, we zoomed into "binary trials", and distinguished S-w from w-S accentuation patterns based on the single-trial β-coefficients from the accent modeling (see methods). The resulting distributions are reported in Fig. 4B, bottom left. Similarly, we disentangled three possible accentuation patterns in the "ternary trials". We performed a separate stepwise regression model using S-w-w, w-S-w, and w-w-S accents as predictors (see methods). Distributions of these accents are reported in Fig. 4B, bottom right. Note that a large proportion of "ternary trials" did not further adhere to one of the ternary accentuation patterns specified as predictors.

This approach allowed showing that, on a portion of trials, individuals spontaneously superimpose accentuation patterns on identical tones embedded in an isochronous equitone sequence. Importantly, the results confirm that the same participants also switched between binary, ternary, and other accentuation patterns over trials. However, in the majority of trials no consistent accentuation pattern was confirmed.

4.5. Binary accents

Once we isolated, at the single-participant level, trials showing binary accentuation patterns, we aimed at statistically testing whether low-β responses would significantly differ in S versus w positions. Thus, we tested whether the modeling approach delivers a meaningful classification of binary accentuation.

We isolated the identified 'binary' accent trials and calculated the tone-by-tone pair-wise difference for low-β across eight positions in the acoustic sequence and preceding the DEV tone. For visualization purposes, the resulting matrix was averaged across trials and the upper symmetrical triangle was masked (Fig. 4D). The original matrix (all trials) was used to calculate metrics of "Binary similarity" and "Binary dissimilarity" (Fig. 4E; see 'Binary accents analyses' in the methods). The Binary similarity features the distributions of amplitude differences on odd- and even-numbered positions. For the "Binary dissimilarity" analyses we calculated the amplitude difference for tones on odd- ver-

sus even-numbered positions (corresponding to accented versus non-accented; thus labeled "Binary difference") and statistically compared it to the Binary similarity (right-side plot in Fig. 4E). Statistical testing yielded a significant difference (FDR-adjusted $p < .05$). Analyses confirmed that the trials classified as 'binary' in the modeling, do indeed show a consistent binary accentuation pattern. Hence, the low-β amplitudes in STD tones in S positions significantly differ from those in w positions. To further verify the validity of the accent modeling approach, we tested whether identified 'preferences for accents' modulate DEV processing.

4.6. DEV processing based on binary accents

We investigated whether DEV processing is modulated by binary accents in 'binary' trials. Thus, we tested whether ERPs to DEV tones falling on S-w positions in the successfully identified "binary" trials would be statistically different. First, we isolated the identified binary trials and discerned 'pure binary' from 'inverse binary' trials based on the beta-coefficient resulting from the regression modeling (see methods). Next, we pooled trials belonging to the same accent (S or w) and statistically compared ERPs to DEV on S-w positions based on the identified accentuation patterns, and thus irrespectively of the sequence position (odd-numbered (9,11th positions) or even-numbered (8,10th)). Within-participant statistical comparison of the respective ERPs yielded significant difference in the time-window between 120–170 ms (*p FDR adjusted <0.05*; Fig. 4F).

Similarly, we tested whether the same S-w effect would be observed for those trials in which no accentuation pattern could be identified ('non-classified' trials). In these non-classified trials, DEV processing was not modulated by binary accentuation patterns (*p FDR adjusted >0.05*; Suppl. Fig. 1). Similarly, DEV processing was not modulated by binary accentuations when pooling all trials (binary, ternary, and non-classified; Suppl. Fig. 1).

Lastly, exploratory analyses focused on the 'ternary' trials and modelled three possible accentuation patterns: S-w-w, w-S-w, w-w-S (Fig. 4b). However, only a small percentage of trials was assigned to these accentuation types (~2% per pattern). This likely reflects that a range of other ternary accents are possible (e.g., 1, −.75, −.25; or 1, −.25, −.75), along with their potential combinations, which were not further modelled here.

In summary, we here show that DEV processing is modulated by binary accents, but exclusively in those trials identified as 'binary' during the accent modeling. This observation supports the modeling procedure as a viable method for identification of trial- and individual-level variability in temporal processing.

5. Discussion

The current study aimed at exploring individual neurophysiological variability in rhythm processing. More specifically, we first examined how delta-band neural activity would track auditory rhythms. Hence, we quantified the sequence-level consistency of phase-alignment towards expected tone onsets. Next, we tested whether neural activity in the low-beta band (12–20 Hz) would reflect the superimposition of binary accentuation patterns, previously described by the "tick-tock" clock phenomenon (Brochard et al., 2003). Accentuations may reflect a neural mechanism which rhythmically and dynamically samples the environment in subunits, resembling the superimposition of a basic beat (strong–weak alternation).

When listening to equitone sequences, participants' neural activity tracked the timing of external events (Fig. 2A), aligning delta-band oscillatory dynamics to expected tone onsets (Fig. 2) (Buzsáki, 2009; Schroeder and Lakatos, 2009; Thut et al., 2012b; Zoefel and VanRullen, 2016). Hence, sequence-level mean vector lengths of delta-band activity preceding tone onsets displayed anticipatory coupling of brain activity to the timing of environmental stimuli. This finding further confirms theoretical views according to which the brain might generate temporal predictions to achieve successful rhythm tracking to optimize sensory processing, perception, and allocation of attention (Friston, 2005; Arnal, 2012; Schröger et al., 2015; Koelsch et al., 2019).

Notably, when pooling phase-angles across sequences either at the single-participant or group-level, we observed random phase distributions. In other words, delta-band neural activity did not always align its high-excitability phase to the expected onset of auditory tones. Rather, we observed a wide range of possible synchronization regimes, which show consistency at the sequence-level, but high variability across sequences and participants. These findings may suggest that rhythm tracking does not necessarily rely on a specific phase-alignment (high-excitability phase) with environmental stimuli to optimize stimulus processing. Rather, our data showed that neural tracking may be supported by an adaptive phase-alignment, that can vary over time and across individuals, while keeping consistency at the trial-level. However, task and attention manipulations may modulate the observed trial-level variability.

Next to rhythm tracking, we investigated the neural signatures associated with the human disposition to accentuate tones when listening to equitone sequences (Brochard et al., 2003). This spontaneous phenomenon typically induces binary (strong-weak (S-w); Brochard et al., 2003) accents, but other accentuation patterns such as ternary ones (S-w-w) are possible (Abecasis et al., 2005; Baath, 2015; Brochard et al., 2003; Fujioka et al., 2012, 2015; Poudrier, 2020). Importantly, while tones are physically identical, these superimposed accents influence observable behavior and underlying neural activity (Nozaradan et al., 2011, 2016, 2017; Exploring How Musical Rhythm Entrains Brain Activity with Electroencephalogram Frequency-Tagging, 2014; Schmidt-Kassow et al., 2011). However, sparse behavioral and neuroimaging research has looked at individual differences (Grahn and Brett, 2007; Grahn and McAuley, 2009) and mainly tested task-based beat processing (e.g., Fujioka et al., 2015). Thus, the question remains whether participants naturally accentuate in the absence of specific task instructions, and if they do so in a consistent manner over time. To address these questions, we focused on participant-level neurophysiological variability and modelled single-trial beta-band activity (Fig. 4A) to probe whether we could observe binary or ternary accentuation patterns. Neural activity in the beta range has been associated with rhythm (Arnal, 2012; Biau and Kotz, 2018; Fujioka et al., 2012, 2015; Morillon et al., 2016) and beat processing (Fujioka et al., 2012, 2010). The current findings confirm its prevalence in time-locked responses to STD tones (Fig. 3B). Furthermore, beta-band activity showed an individuals' spontaneous disposition to superimpose accentuation patterns, even when not instructed to do so (Fig. 4). Hence, we characterized inter- and within-participant differences in adopting binary and ternary accents over time (Fig. 4B and Suppl. Table 2). These accentuations might reflect the automatic predisposition to sample continuous auditory input streams into predictable, coherent, and finite units. This might reflect fluctuations of attentional resources over time (Bolton, 1984; Jones, 1976; Schroeder and Lakatos, 2009) (rather than being equally distributed over time), and variations within each attentional cycle so to attribute salience to accentuated events. We tested this view by particularly focusing on binary accents (Fig. 4C-F) and showed that neural responses in the beta-band for tones falling on odd-numbered ('strong' ('S')) positions were significantly greater from those in even-numbered positions ('weak' ('w'); Fig. 4E, F) on a selection of trials. This suggests that individuals spontaneously superimpose binary accents (S-w) while listening to equitone sequences to parse and segment continuous sensory streams, potentially allocating attentional resources to salient sensory events and to optimize perception (Nobre and Van Ede, 2018; Shalev et al., 2019). However, most trials (~60%) did not reflect either binary or ternary accents (both approximately in 20% of trials, Fig. 4B). This observation may result for at least four reasons: (i) individuals do not always accentuate, when not instructed to do so; (ii) individuals do accentuate, but switch between accentuation patterns over time; (iii) the applied method or the

data per se are not sensitive enough to pick up spontaneous and varied accentuations; (iv) the employed stimulation rate may be suboptimal to induce different accentuation patterns such as ternary and longer patterns (Baath, 2015; Poudrier, 2020). Indeed, prior studies on predictive timing and beta-band actvity relied on source-reconstructed neural MEG data (Fujioka et al., 2012, 2015), thus probably benefitted from a higher signal-to-noise ratio than the current study. Furthermore, although it is known cognitive bias might induce the structuring of tonal sequences into groups of two (Polak et al., 2018; Savage et al., 2015) rather than other groupings, other studies have shown a link between stimulation rates and preferred accentuation patterns (Baath, 2015; Poudrier, 2020). Accordingly, inter-stimulus intervals between 500–900 ms should preferentially induce a binary accent (and its double, i.e., accents over groups of four tones), while ternary and other accentuations are possible, although less common, at this and slower rates.

In the absence of perceptual reports, we cannot preclude these options. However, we show positive evidence that in a portion of trials participants display clear binary and ternary accentuation patterns independent of behavioral reporting or finger or foot tapping. These results support the notion that humans sample the environment in an individual, dynamic manner (Large and Jones, 1999).

We note that classifying trials based on beta fluctuations, followed by testing beta time-locked responses over sequence positions might be circular, but is confirmatory. Thus, we compared neural responses to unpredicted (deviant) tones falling on S-w positions in the classified binary trials and found a significant effect in ERP responses to DEV tones (Fig. 4F). Notably, the modulation of DEV processing was absent in those trials which did not adhere to a specific accentuation pattern (non-classified trials; Suppl. Fig. 1). These observations parallel earlier findings (Brochard et al., 2003; Jongsma et al., 2004; Schmidt-Kassow et al., 2011) and indicate that accentuations might affect how we allocate attention to the auditory environment. However, further investigations are needed to clarify this intricate link between attention deployment, attentional shifts, and temporal processing in listening contexts. Thus, future studies could further dive in the frequency-specificity of such effects and complement neural data with perceptual reports.

In summary, we characterize individual neurophysiological signatures of temporal processing, and associate them with specific delta-band phase-coupling mechanisms and with beta-band dynamics, respectively. The findings showcase the feasibility of using EEG to identify individual neurophysiological signatures of temporal processing, suggesting that common trial- and group-level averaging approaches might inevitably obscure inter-individual differences and trial-by-trial variability. In contrast, the approach adopted here allows the monitoring of neurophysiological variability underlying *flexible but consistent mechanisms* for evaluating and adapting to (un)predictable environmental stimuli. Consequently, we propose that zooming into individual variability might allow to better predict behavioral variability in processing simple and complex environmental rhythms (e.g., speech tracking; Kandylaki and Criscuolo, 2021) in neurotypical and pathological populations (Schwartze et al., 2015, 2016).

6. Conclusions

When listening to isochronous equitone sequences, humans' neural activity tends to spontaneously align and track the timing of auditory events. A novel trial-level modeling approach additionally confirms that individuals tend to superimpose accentuation patterns onto isochronous equitone sequences, indicating active sampling of the acoustic environment and (potentially) differential allocation of cognitive resources. We explored inter-individual and trial-level neurophysiological variability in temporal processing and auditory accentuation, and reveal flexible, time-varying neural mechanisms involved in effective evaluation and adaptation to environment rhythms. The combined findings highlight that an individualized analysis approach to neurophysiological data can indicate meaningful variation in a listening context and should be considered in a more differentiated account of the role of temporal dynamics in audition.

Credit author statement

SAK and CO conceptualized the study.
CO collected the data.
AC, SAK and MS processed the data.
AC, MS, MJH, CO and SAK interpreted the results and wrote the manuscript.

Declaration of Competing Interest

None to declare.

Data availability

Data will be made available on request.

Acknowledgments

We thank Ina Koch at the Max Planck for Human Cognitive and Brain Sciences, Dept. of Neuropsychology, Leipzig, Germany for her support in data collection. This work is supported by a grant of the Dutch Research Council (NWO - 406.18.GO.063) and the Van der Gaag Fund, Royal Netherlands Academy of Arts & Sciences.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.neuroimage.2023.120090.

References

Abbasi, O., Gross, J., 2020. Beta-band oscillations play an essential role in motor–auditory interactions. Hum. Brain Mapp. 41 (3), 656–665. doi:10.1002/hbm.24830.
Abecasis, D., Brochard, R., Granot, R., Drake, C., 2005. Differential brain response to metrical accents in isochronous auditory sequences. Music Percept. doi:10.1525/mp.2005.22.3.549.
Arnal, L.H., 2012. Predicting "When" using the motor system's beta-band oscillations. Front. Hum. Neurosci. 6. doi:10.3389/fnhum.2012.00225.
Baath, R., 2015. Subjective rhythmization: a replication and an assessment of two theoretical explanations. Music Percept. 33 (2), 244–254.
Barnes, R., Jones, M.R., 2000. Expectancy, attention, and time. Cogn. Psychol. doi:10.1006/cogp.2000.0738.
Bartolo, R., Merchant, H., 2015. β oscillations are linked to the initiation of sensory-cued movement sequences and the internal guidance of regular tapping in the monkey. J. Neurosci. doi:10.1523/JNEUROSCI.4570-14.2015.
Berens, P., 2009. CircStat: a MATLAB toolbox for circular statistics. J. Stat. Softw. 31 (10), 1–21. doi:10.18637/JSS.V031.I10.
Biau, E., Kotz, S.A., 2018. Lower beta: a central coordinator of temporal prediction in multimodal speech. Front. Hum. Neurosci. 12 (October), 1–12. doi:10.3389/fnhum.2018.00434.
Bolton, 1984. Rhythm. Am. J. Psychol. 6 (2), 1–46.
Brochard, R., Abecasis, D., Potter, D., Ragot, R., Drake, C., 2003. The "Ticktock" of our internal clock. Psychol. Sci. 14 (4), 362–366. doi:10.1111/1467-9280.24441.
Buzsáki, G. (2009). Rhythms of the brain. In *Rhythms of the Brain*. 10.1093/acprof:oso/9780195301069.001.0001
Cohen, M.X., 2014. Analyzing Neural Time Series Data: Theory and Practice. MIT Press doi:10.1017/CBO9781107415324.004.
Colling, L.J., Noble, H.L., Goswami, U., 2017. Neural entrainment and sensorimotor synchronization to the beat in children with developmental dyslexia: an EEG study. Front. Neurosci. 11 (JUL). doi:10.3389/fnins.2017.00360.
Ding, N., Melloni, L., Zhang, H., Tian, X., Poeppel, D., 2015. Cortical tracking of hierarchical linguistic structures in connected speech. Nat. Neurosci. 19 (1), 158–164. doi:10.1038/nn.4186.
Doelling, K.B., Poeppel, D., 2015. Cortical entrainment to music and its modulation by expertise. Proc. Natl. Acad. Sci. U.S.A. 112 (45), E6233–E6242. doi:10.1073/pnas.1508431112.
Friston, K., 2005. A theory of cortical responses. Philos. Trans. R. Soc. B Biol. Sci. doi:10.1098/rstb.2005.1622.
Fujioka, T., Ross, B., Trainor, L.J., 2015. Beta-band oscillations represent auditory beat and its metrical hierarchy in perception and imagery. J. Neurosci. doi:10.1523/JNEUROSCI.2397-15.2015.

Fujioka, T., Trainor, L.J., Large, E.W., Ross, B., 2012. Internalized timing of isochronous sounds is represented in neuromagnetic beta oscillations. J. Neurosci. 32 (5), 1791–1802. doi:10.1523/JNEUROSCI.4107-11.2012.

Fujioka, T., Zendel, B.R., Ross, B., 2010. Endogenous neuromagnetic activity for mental hierarchy of timing. J. Neurosci. 30 (9), 3458–3466. doi:10.1523/JNEUROSCI.3086-09.2010.

Giraud, A.L., Poeppel, D., 2012. Cortical oscillations and speech processing: emerging computational principles and operations. Nat. Neurosci. 15 (4), 511–517. doi:10.1038/nn.3063.

Goswami, U., 2011. A temporal sampling framework for developmental dyslexia. Trends Cogn. Sci. 15 (1), 3–10. doi:10.1016/j.tics.2010.10.001, *(Regul. Ed.)*.

Grahn, J.A., Brett, M., 2007. Rhythm and beat perception in motor areas of the brain. J. Cogn. Neurosci. 19 (5), 893–906. doi:10.1162/jocn.2007.19.5.893.

Grahn, J.A., McAuley, J.D., 2009. Neural bases of individual differences in beat perception. Neuroimage 47 (4), 1894–1903. doi:10.1016/J.NEUROIMAGE.2009.04.039.

Haegens, S., Zion Golumbic, E., 2018. Rhythmic facilitation of sensory processing: a critical review. Neurosci. Biobehav. Rev. 86 (November 2017), 150–165. doi:10.1016/j.neubiorev.2017.12.002.

Jones, M.R., 1976. Time, our lost dimension: toward a new theory of perception, attention, and memory. Psychol. Rev. 83 (5), 323–355. doi:10.1037/0033-295X.83.5.323.

Jongsma, M.L.A., Desain, P., Honing, H., 2004. Rhythmic context influences the auditory evoked potentials of musicians and non-musicians. Biol. Psychol. 66 (2), 129–152. doi:10.1016/J.BIOPSYCHO.2003.10.002.

Kandylaki, K.D., Criscuolo, A., 2021. Neural tracking of speech: top-down and bottom-up influences in the Musician's brain. J. Neurosci. 41 (31), 6579–6581. doi:10.1523/JNEUROSCI.0756-21.2021.

Kaneshiro, B., Nguyen, D.T., Norcia, A.M., Dmochowski, J.P., Berger, J., 2020. Natural music evokes correlated EEG responses reflecting temporal structure and beat. Neuroimage 214, 116559. doi:10.1016/J.NEUROIMAGE.2020.116559.

Koelsch, S., Vuust, P., Friston, K., 2019. Predictive processes and the peculiar case of music. Trends Cogn. Sci. doi:10.1016/j.tics.2018.10.006, *(Regul. Ed.)*.

Kononowicz, T.W., van Rijn, H., 2015. Single trial beta oscillations index time estimation. Neuropsychologia 75, 381–389. doi:10.1016/J.NEUROPSYCHOLOGIA.2015.06.014.

Kotz, S.A., Schwartze, M., 2010. Cortical speech processing unplugged: a timely subcortico-cortical framework. Trends Cogn. Sci. 14 (9), 392–399. doi:10.1016/j.tics.2010.06.005, *(Regul. Ed.)*.

Lakatos, P., Karmos, G., Mehta, A.D., Ulbert, I., Schroeder, C.E., 2008. Entrainment of neuronal oscillations as a mechanism of attentional selection. Science 320 (5872), 110–113. doi:10.1126/science.1154735.

Large, E.W., Jones, M.R., 1999. The dynamics of attending: how people track time-varying events. Psychol. Rev. 106 (1), 119–159. doi:10.1037/0033-295X.106.1.119.

Merchant, H., Bartolo, R., 2018. Primate beta oscillations and rhythmic behaviors. In: Journal of Neural Transmission, 125. Springer-Verlag Wien, pp. 461–470. doi:10.1007/s00702-017-1716-9.

Merchant, H., Grahn, J., Trainor, L., Rohrmeier, M., Fitch, W.T., 2015. Finding the beat: a neural perspective across humans and non-human primates. Philos. Trans. R. Soc. B Biol. Sci. doi:10.1098/rstb.2014.0093.

Merker, B.H., Madison, G.S., Eckerdal, P., 2009. On the role and origin of isochrony in human rhythmic entrainment. Cortex doi:10.1016/j.cortex.2008.06.011.

Møller, C., Stupacher, J., Celma-Miralles, A., Vuust, P., 2021. Beat perception in polyrhythms: time is structured in binary units. PLoS One 16 (8), e0252174. doi:10.1371/JOURNAL.PONE.0252174.

Morillon, B., Schroeder, C.E., Wyart, V., Arnal, L.H., 2016. Temporal prediction in lieu of periodic stimulation. J. Neurosci. doi:10.1523/JNEUROSCI.0836-15.2016.

Mormann, F., Fell, J., Axmacher, N., Weber, B., Lehnertz, K., Elger, C.E., Fernández, G., 2005. Phase/amplitude reset and theta-gamma interaction in the human medial temporal lobe during a continuous word recognition memory task. Hippocampus 15 (7), 890–900. doi:10.1002/hipo.20117.

Nave, K.M., Hannon, E.E., Snyder, J.S., 2022. Steady state-evoked potentials of subjective beat perception in musical rhythms. Psychophysiology 59 (2), e13963. doi:10.1111/PSYP.13963.

Nobre, Anna C., Van Ede, F., 2018. Anticipated moments: temporal structure in attention. Nat. Rev. Neurosci. 19 (1), 34–48. doi:10.1038/nrn.2017.141.

Exploring how musical rhythm entrains brain activity with electroencephalogram frequency-tagging, 369 Philos. Trans. R. Soc. B Biol. Sci. (2014). doi:10.1098/rstb.2013.0393.

Nozaradan, S., Peretz, I., Keller, P.E., 2016. Individual differences in rhythmic cortical entrainment correlate with predictive behavior in sensorimotor synchronization. Sci. Rep. doi:10.1038/srep20612.

Nozaradan, S., Peretz, I., Missal, M., Mouraux, A., 2011. Tagging the neuronal entrainment to beat and meter. J. Neurosci. 31 (28), 10234–10240. doi:10.1523/JNEUROSCI.0411-11.2011.

Nozaradan, S., Schwartze, M., Obermeier, C., Kotz, S.A., 2017. Specific contributions of basal ganglia and cerebellum to the neural tracking of rhythm. Cortex 95, 156–168. doi:10.1016/j.cortex.2017.08.015.

Nozaradan, S., Zerouali, Y., Peretz, I., Mouraux, A., 2015. Capturing with EEG the neural entrainment and coupling underlying sensorimotor synchronization to the beat. Cereb. Cortex 25 (3), 736–747. doi:10.1093/cercor/bht261.

Obleser, J., Herrmann, B., Henry, M.J., 2012. Neural oscillations in speech: don't be enslaved by the envelope. Front. Hum. Neurosci. 0 (AUGUST), 250. doi:10.3389/fnhum.2012.00250.

Oostenveld, R., Fries, P., Maris, E., Schoffelen, J.M., 2011. FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. Comput. Intell. Neurosci. doi:10.1155/2011/156869.

Patel, A.D., Iversen, J.R., 2014. The evolutionary neuroscience of musical beat perception: the Action Simulation for Auditory Prediction (ASAP) hypothesis. Front. Syst. Neurosci. doi:10.3389/fnsys.2014.00057.

Polak, R., Jacoby, N., Fischinger, T., Goldberg, D., Holzapfel, A., London, J., 2018. Rhythmic prototypes across CulturesA comparative study of tapping synchronization. Music Percept. 36 (1), 1–23. doi:10.1525/MP.2018.36.1.1.

Poudrier, È., 2020. The influence of rate and accentuation on subjective rhythmization. Music Percept. 38 (1), 27–45. doi:10.1525/mp.2020.38.1.27.

Savage, P.E., Brown, S., Sakai, E., Currie, T.E., 2015. Statistical universals reveal the structures and functions of human music. Proc. Natl. Acad. Sci. U.S.A. 112 (29), 8987–8992. doi:10.1073/PNAS.1414495112/SUPPL_FILE/PNAS.1414495112.SD01.XLS.

Schmidt-Kassow, M., Rothermich, K., Schwartze, M., Kotz, S.A., 2011. Did you get the beat? Late proficient French-German learners extract strong–weak patterns in tonal but not in linguistic sequences. Neuroimage 54 (1), 568–576. doi:10.1016/j.neuroimage.2010.07.062.

Schroeder, C.E., Lakatos, P., 2009. Low-frequency neuronal oscillations as instruments of sensory selection. Trends Neurosci. 32 (1), 9–18. doi:10.1016/j.tins.2008.09.012.

Schröger, E., Kotz, S.A., SanMiguel, I., 2015. Bridging prediction and attention in current research on perception and action. Brain Res. 1626, 1–13. doi:10.1016/j.brainres.2015.08.037.

Schwartze, M., Keller, P.E., Kotz, S.A., 2016. Spontaneous, synchronized, and corrective timing behavior in cerebellar lesion patients. Behav. Brain Res. 312, 285–293. doi:10.1016/j.bbr.2016.06.040.

Schwartze, M., Rothermich, K., Schmidt-Kassow, M., Kotz, S.A., 2011. Temporal regularity effects on pre-attentive and attentive processing of deviance. Biol. Psychol. 87 (1), 146–151. doi:10.1016/j.biopsycho.2011.02.021.

Schwartze, M., Stockert, A., Kotz, S.A., 2015. Striatal contributions to sensory timing: voxel-based lesion mapping of electrophysiological markers. Cortex 71, 332–340. doi:10.1016/j.cortex.2015.07.016.

Shalev, N., Nobre, K., & van Ede, F. (2019). *Time for what: breaking down temporal anticipation*.

Thut, G., Miniussi, C., Gross, J., 2012. The functional importance of rhythmic activity in the brain. Curr. Biol. 22 (16), R658–R663. doi:10.1016/j.cub.2012.06.061.

Waschke, L., Kloosterman, N.A., Obleser, J., Garrett, D.D., 2021. Behavior needs neural variability. Neuron 109 (5), 751–766. doi:10.1016/J.NEURON.2021.01.023.

Zoefel, B., ten Oever, S., Sack, A.T., 2018. The involvement of endogenous neural oscillations in the processing of rhythmic input: more than a regular repetition of evoked neural responses. Front. Neurosci. 12 (MAR), 1–13. doi:10.3389/fnins.2018.00095.

Zoefel, B., VanRullen, R., 2016. EEG oscillations entrain their phase to high-level features of speech sound. Neuroimage 124, 16–23. doi:10.1016/j.neuroimage.2015.08.054.

Further reading

Háden, G.P., Honing, H., Török, M., Winkler, I., 2015. Detecting the temporal structure of sound sequences in newborn infants. Int. J. Psychophysiol. doi:10.1016/j.ijpsycho.2015.02.024.

Nobre, A.C., Rohenkohl, G., Stokes, M., 2012. Nervous anticipation : top-down biasing across space and time. Cogn. Neurosci. Atten..