Title:

Neural entrainment predicts the anticipatory P300 component during musical meter perception: An EEG study using dual-meter sound stimuli

Abbreviated title:

Neural entrainment predicts P300 during meter perception

Authors:

Sotaro Kondoh[1,2,3,4,5]

Kazuo Okanoya[1,2&]

Ryosuke O. Tachibana[1,2,6*]

[1] Graduate School of Arts and Sciences, The University of Tokyo, Tokyo, Japan

[2] Advanced Comprehensive Research Organization, Teikyo University, Tokyo, Japan

[3] Graduate School of Media and Governance, Keio University, Fujisawa, Japan

[4] Faculty of Environment and Information Studies, Keio University, Fujisawa, Japan

[5] Japan Society for the Promotion of Science, Tokyo, Japan

[6] Human Informatics and Interaction Research Institute, National Institute of Advanced Industrial Science and Technology, Tsukuba, Japan

[&] Co-corresponding author

E-mail: kazuookanoya@gmail.com

[*] Corresponding author

E-mail: rtachi@gmail.com

Abstract

Meter, the organization of beats into regular groupings, is a core element of music perception. Switches in meters frequently occur in music, reflecting the interplay between top-down attentional processes and the bottom-up processing of acoustic features. To disentangle this interplay in meter perception, we had previously invented a sound stimulus that simultaneously contained two meters. This dual-meter stimulus consisted of band-limited noise bursts that varied in center frequency and duration, with one feature following a triple-meter pattern and the other following a quadruple-meter pattern. Thus, the perceived meter switches between the two by changing attention from one acoustic feature to another, even when listening to identical stimuli. Using these stimuli, we here demonstrated the neural processes involved in meter perception by recording electroencephalogram (EEG) signals, while dissociating the influence of acoustic differences. We first found that the perceived meter structure was reflected in the entrained neural oscillations by comparing the spectral profiles of EEG signals when the participants focused on triple-meter versus quadruple-meter features. Second, an event-related potential component for anticipation (P300) was observed when the expected meter structure was disrupted by altering the frequency or duration of one sound at the end of the stimulus. Notably, individuals with stronger neural entrainment exhibited larger P300 responses to these disruptions. Our findings suggest that top-down attention modulates anticipation of the meter structure, and this anticipation induces neural entrainment that is involved in meter perception.

***Keywords**:*
meter, dual-meter stimuli, electroencephalogram, neural entrainment, P300

Significance Statement

Perceiving meters is fundamental to music cognition, integrating the bottom-up processing of acoustic features with top-down attention directed toward those features. Traditional methods have struggled to separate attention-driven neural entrainment and anticipation related to meter perception from acoustic variability. Here, we introduce unique “dual-meter” stimuli that listeners perceive as different meters when attending to frequency and duration. EEG recordings with these stimuli revealed that neural synchronization to perceived meters predicts the anticipatory P300 response of the brain to metrical violations. These results demonstrate that feature-specific attention modulates both entrainment and expectation, providing new insights into the dynamic neural processes that underlie meter perception.

Introduction

Meter is one of the core elements of music perception that emerges from its ability to extract regular beats and assign varying levels of subjective intensity to them, thereby forming organized groupings (Cooper & Meyer, 1963; Fujii & Schlaug, 2013; Honing, 2013; Kotz et al., 2018; Large & Jones, 1999; Lerdahl & Jackendoff, 1983; London, 2012; Palmer & Krumhansl, 1990; Patel, 2006, 2010). In a duple meter found in march music, beats follow a recurring "strong(s)-weak(w)-s-w-…" pattern. The triple meter organizes beats as "s-w-w-s-w-w-…," which typically characterizes waltz music.

Meter perception involves both the bottom-up processing of acoustic features and top-down attentional mechanisms. A sound sequence such as "loud-soft-loud-soft…" is perceived as a duple meter, with louder sounds marking strong beats (Chen et al., 2006, 2009; Snyder & Large, 2005; Zatorre et al., 2007), elucidating the involvement of the bottom-up process. On the other hand, because music normally contains a variety of acoustic cues, actual meter perception should also require top-down selective attention to specific features. This is analogous to the "cocktail party effect" and auditory stream segregation, where listeners focus on a particular auditory stream among competing ones (Bregman, 1994; Bregman et al., 2001; Cherry, 1953). In our previous study (Kondoh et al., 2021), we tested this issue with "dual-meter" stimuli that allowed us to perceive two meters, which were created from sequences of band-limited noise bursts with varying center frequencies and durations. For instance, one of the stimuli had a frequency pattern "high-low-low…" to establish a triple meter while a duration pattern of "long-short-short-short…" corresponded to a quadruple meter at the same time. By instructing participants to shift their attention between frequency and duration, we found that listeners experienced different meters despite the physical sound sequences remaining intact. This result demonstrates that dual-meter stimuli offer a novel experimental paradigm for investigating the neural mechanisms of meter perception while dissociating the influence of acoustic variability.

Recent research suggests that neural entrainment and anticipatory processing are the key mechanisms underlying meter perception (Harding et al., 2025; Snyder et al., 2024; Vuust et al., 2022). Neural entrainment refers to the automatic synchronization of endogenous brain rhythms with external auditory inputs (Fujioka et al., 2009; Nozaradan et al., 2011), supporting bottom-up processing. At the same time, these oscillatory neural activities can shape the anticipation of future events according to dynamic attending theory (Ellis & Jones, 2009; Jones, 2018; Jones & Boltz, 1989). Thus, successful top-down selective attention to a particular feature of music may promote both neural synchronization

and anticipation specific to that feature, thereby forming a perception of the meter. Nonetheless, empirical evidence for this mechanism remains unclear because of methodological constraints in isolating neural activity related to entrainment and anticipation from acoustic influence.

To address this gap, this study employed dual-meter stimuli in an EEG experiment. First, we compared the spectral power when participants attended to either frequency or duration, allowing us to assess neural entrainment to the perceived meter while minimizing the confounding effects of acoustic differences. Second, we analyzed event-related potentials (ERPs) elicited by violations of metrical expectations introduced by altering the frequency or duration of a single noise burst. Such violations are known to evoke mismatch negativity (MMN), which reflects the automatic detection of prediction errors, and P300, which is associated with the conscious evaluation of unexpected events (Koelsch et al., 2019; Näätänen et al., 1978; Sutton et al., 1965; Vuust et al., 2009). By comparing ERP amplitudes across attentional conditions, we aimed to identify components linked to anticipatory processing, rather than acoustic changes. Finally, we examined whether neural entrainment indices predicted the MMN and P300 amplitudes. We hypothesized that stronger neural entrainment would be associated with larger ERP responses, indicating a functional relationship between entrainment and anticipation.

Materials and Methods

Participants

Thirty-four participants (17 males and 17 females; mean age = 19.8 ± 1.6 years) were recruited for this study from March 9 to October 1, 2022. All participants reported normal hearing and no history of hearing impairments. Before participating, they received instructions regarding the study and signed a written informed consent form. This study was approved by the Ethical Review Committee for Experimental Research on Human Subjects at the Graduate School of Arts and Sciences, the University of Tokyo, Japan (#663). The sample size was determined a priori using G*Power 3 (Faul et al., 2007) to ensure sufficient power (0.80) for detecting medium effect sizes ($\alpha = 0.05$). The analysis focused on pairwise comparisons of neural entrainment indices, specifically examining spectral differences across meter-related frequency bands (see ***Statistics*** below).

Stimuli

We designed dual-meter stimuli (see **Fig. 1A, B**) consisting of 24 band-noise bursts. Each burst had one of two center frequencies: a high frequency of 1600 Hz or a low frequency of 1400 Hz, resulting

in a frequency difference of 200 Hz. We also varied the duration of the bursts and classified them as long (150 ms) or short (50 ms) sounds, creating a duration difference of 100 ms. These specific frequency and duration differences were selected based on our previous study (Kondoh et al., 2021), which demonstrated that listeners were most likely to switch attention between acoustic features under these parameters. The inter-onset intervals (IOIs) were fixed at 300 ms. The bandwidth of the noise bursts was 120% of the center frequency, with rise and decay times of 10 ms.

We created two types of stimuli based on regularity patterns in center frequency and duration: frequency-triple stimuli (see **Fig. 1A**), and frequency-quadruple stimuli (see **Fig. 1B**). Participants were randomly assigned to one of these two stimulus categories, with half listening to the frequency-triple stimuli and the other half listening to the frequency-quadruple stimuli. In the frequency-triple condition, the frequency pattern repeated every three sounds ("high–low–low" or "low–high–high"), while the duration pattern repeated every four sounds ("long–short–short–short" or "short–long–long–long"). In contrast, the frequency-quadruple category featured a frequency pattern that repeated every four sounds ("high–low–low–low" or "low–high–high–high"), paired with a duration pattern that repeated every three sounds ("long–short–short" or "short–long–long"). Consequently, each category contained four stimulus subtypes determined by which sounds served as strong beats: high and long (see **Fig. 1A, B**), high and short (see **Fig. S1A, D**), low and long (see **Fig. S1B, E**), and low and short (see **Fig. S1C, F**).

Procedure

Pre-training to familiarize with dual-meter stimuli

Because the participants were unfamiliar with dual-meter stimuli, we conducted an online pre-training session to help them focus on center frequency and duration differences. The training program was developed using lab.js (Henninger et al., 2022), and the participants completed the program individually on their personal computers using headphones. To verify the use of headphones, we used a screening test (Woods et al., 2017). During the task, participants listened to a dual-meter stimulus and were instructed to count specific sounds categorized as "high," "low," "long," or "short." Each stimulus contained one to eight target sounds, resulting in a total of 3 (1 × 3: triple) to 32 (8 × 4: quadruple) sounds per trial. After listening, they selected numbers from 1 to 8 on the screen that corresponded to the count of the target sounds.

The training consisted of six blocks of 32 trials (4 characteristics × 8 targets). In the high/low counting

trials, the frequency difference remained at 200 Hz, whereas the duration difference increased from 64 ms in the first and second blocks to 80 ms in the third and fourth blocks, rendering the frequency difference less prominent. Conversely, in the long/short counting trials, the duration difference remained at 100 ms, whereas the frequency difference increased from 128 Hz in the first and second blocks to 160 Hz in the third and fourth blocks, making the duration difference less pronounced. In the final two blocks, both tasks used stimuli with a frequency difference of 200 Hz and a duration difference of 100 ms, matching the parameters used in the EEG recording. No feedback on the correct answers was provided during the training.

EEG recording

EEG recordings were acquired using a 32-channel EasyCap and BrainAmp DC system (Brain Products GmbH, Germany) with a sampling rate of 1000 Hz. The electrodes were positioned at C3, C4, CP1, CP2, CP5, CP6, Cz, F3, F4, F7, F8, FC1, FC2, FC5, FC6, FT10, Fp1, Fp2, Fz, Iz, O1, O2, P3, P4, P7, P8, PO10, PO9, Pz, T7, and T8. In addition, an electrode (AFp10) was placed near the right eye to record the electrooculogram activity, and reference electrodes were attached to the left (A1) and right (A2) earlobes.

During the recording session, the participants listened to the dual-meter stimuli while attending to differences in either frequency or duration (see **Fig. 1C**). EEG data from the first 18 noise bursts were analyzed to assess neural entrainment indicative of the musical meter. To disrupt metrical anticipation, the center frequency or duration of the 19th, 20th, or 21st band-noise bursts was modified (referred to as altered sounds; see **Fig. S2**). For ERP analysis, standard sounds were defined as those with both center frequency and duration matching the altered sounds. Two conditions were distinguished: the violated condition, where changes in the attended acoustic feature disrupted metrical anticipation, and the unviolated condition, where the metrical structure remained intact despite modifications.

Each trial began with 0.8 seconds of silence followed by a 7.2-second stimulus presentation. A block consisted of 14 trials that focused on one acoustic feature, including six trials under the violated condition, six under the unviolated condition, and two trials in which the stimulus was presented without any modifications. Throughout each block, a fixation cross and a display indicating the target feature were shown on the screen. Each session comprised four blocks with randomized stimulus assignments and trial orders, and the participants completed a total of four sessions.

Experimental control was achieved using Presentation 18.1 software (Neurobehavioral Systems, Inc., USA). Instructions about the attended acoustic feature and the fixation cross were presented on a 24-inch LCD monitor (XL2420T, BenQ, Taiwan), with participants seated approximately 80 cm from the screen. Auditory stimuli were delivered binaurally through earphones (Compumedics Neuroscan Inc., USA) at approximately 70 dB SPL via an external sound card (UA-22, Roland, Japan).

Analysis

Pre-processing of EEG

EEG preprocessing was conducted using MATLAB R2023b and EEGLAB v2022.1 (Delorme & Makeig, 2004). For each participant, the data were filtered with a high-pass filter set at 0.5 Hz and a low-pass filter set at 30 Hz. Electrodes with impedance values greater than 15 kΩ were excluded from further analysis. Epochs were extracted for each trial, yielding 224 epochs (28 stimuli × 2 attended acoustic features × 4 sessions). An independent component analysis (ICA) was then performed to decompose the EEG data into a number of components equal to the number of electrodes (Delorme et al., 2007; Makeig et al., 2004). The ICLabel tool was used to classify these components as either brain-derived signals or noise sources, including ocular, muscular, and cardiac activities, as well as line noise (Pion-Tonachini et al., 2019). For subsequent analyses, only the components exhibiting more than 50% brain-derived signals were retained.

Neural entrainment related to meter

First, we identified the most appropriate electrodes for the analysis of neural oscillations. We extracted EEG data corresponding to the first 18 noise bursts of the dual-meter stimuli for each participant and electrode (totaling 5.4 seconds, 0.3 seconds × 18 bursts). Trials were pooled without distinguishing whether the participants attended to triple- or quadruple-related features. To remove slow drifts, a moving median filter with a 1.2-second window was applied. The EEG signal was then tapered using a raised cosine window that gradually increased over the first 0.9 seconds and decreased over the last 0.3 seconds. The data were resampled to 40 Hz for computational efficiency and then zero-padded to 1,728 points (5.4 seconds × 40 Hz × 8) for improved frequency resolution, yielding a resolution of approximately 0.023 Hz (40 Hz / 1,728 points).

After applying a Fast Fourier Transform (FFT), the spectra for each electrode were averaged across participants (see **Fig. 2A**). To establish a baseline, we generated a surrogate signal by rearranging the original EEG data every 300 ms, matching the inter-onset intervals of the noise bursts. This surrogate

signal underwent the same raised cosine tapering and FFT processing, yielding the same frequency resolution (see **Fig. 2A**). For each electrode, we computed the amplitude ratio by dividing the original signal by the surrogate signal, converting the result into logarithms, and then averaging across participants. We focused on the frequency of approximately 3.33 Hz (ranging from 2.83 to 3.83 Hz), which reflects neural synchronization to every noise burst, and compared the magnitude of this ratio across electrodes. The electrode with the highest ratio, along with its surrounding electrodes, was used for neural entrainment analysis (see **Fig. 2B**).

We then calculated the neural indices reflecting meter-related entrainment. Previous studies used the amplitude of EEG potentials at frequencies corresponding to the perceived meter, known as steady-state evoked potentials (SSEPs), as an index of neural entrainment (Chemin et al., 2014; Cirelli et al., 2016; Gibbings et al., 2023; Nave et al., 2022; Nozaradan et al., 2011, 2012, 2016, 2018). However, calculating SSEPs typically involves subtracting the average amplitude of two neighboring frequency bins from each target bin after FFT, using neighboring frequencies as a baseline for the target frequency. This approach can limit the accuracy of the comparisons across different frequency bins. To address this limitation, we employed a novel method by calculating the spectral difference in EEG signals between states in which participants attended to triple-related versus quadruple-related features. Because the stimulus was physically identical regardless of the attended feature, subtraction of the spectra was assumed to eliminate acoustic variability and capture meter-related neural activity.

We defined five frequency bands corresponding to the number of beats in each group (see **Fig. 2C**): Band-1, centered at 3.33 Hz (range 3.26–3.43 Hz), corresponding to a single beat; Band-2, centered at 1.67 Hz (range 1.60–1.76 Hz), corresponding to two beats; Band-3, centered at 1.11 Hz (range 1.04–1.20 Hz), corresponding to a triple meter; Band-4, centered at 0.83 Hz (range 0.76–0.93 Hz), corresponding to a quadruple meter; and Band-5, centered at 0.67 Hz (range 0.60–0.76 Hz), corresponding to five beats. We computed the spectral amplitude for each band when participants attended to triple-related features and quadruple-related features, and then calculated the spectral difference by subtracting the quadruple-related amplitude from the triple-related amplitude. A more positive value indicates stronger neural entrainment to the triple feature, whereas a more negative value indicates stronger entrainment to the quadruple feature (see **Fig. 2D**).

Mismatch negativity and P300 components

Electrodes were selected to analyze mismatch negativity (MMN) and P300 components associated

with metrical anticipation. ERPs for both standard and altered sounds were computed by averaging the data across all participants without differentiating between violated and unviolated conditions (see **Fig. 3A**). The analysis window ranged from 0 ms (sound onset) to 600 ms, with baseline correction applied using the 100 ms preceding sound onset. The MMN latency was determined by identifying the negative deflection in the amplitude difference between the altered and standard sounds. The electrode exhibiting the largest negative deflection, along with its neighboring electrodes, was chosen for further MMN analysis (see **Fig. 3B**). As the P300 typically follows the MMN, the same electrodes were used for the analysis, with the latency range beginning after the MMN period and extending to 600 ms.

ERPs were then obtained by averaging the signals for both the altered and standard sounds under each condition (see **Fig. 3C, D**). The analysis focused on a latency window from 0 ms at sound onset to 600 ms using the 100 ms preceding sound onset as the baseline. We subtracted the ERP elicited by standard sounds from that elicited by altered sounds (see **Fig. 3E**). This amplitude difference allowed us to assess whether the MMN and P300 components reflected violations of the anticipatory metrical structure induced by the altered sounds, while controlling for acoustic variability.

Statistics

First, we tested whether the spectral differences varied among meter-related frequency bands. Because the spectral difference represents the difference in spectra between situations in which participants attended to triple- versus quadruple-related acoustic features for an identical stimulus, this measure should remain constant across frequency bands if it does not reflect meter-specific neural synchronization. The data were fitted using a linear mixed-effects regression model, with spectral difference as the dependent variable, meter-related frequency band as a fixed effect, and participant ID as a random effect. We used the 'lmer' function in the 'lme4' package (vthe er. 1.1.35.1) (Bates et al., 2015), with significance tests were conducted by 'lmerTest' package (ver. 3.1.3) (Kuznetsova et al., 2017). After fitting the model to the data, we conducted pairwise comparisons between meter-related frequency bands using the 'emmeans' package (ver. 1.10.3) (Lenth, 2024). *P*-values of multiple comparisons were adjusted by Tukey-Kramer method (denoted as p_{Tukey}).

Next, we investigated whether the amplitude difference (altered minus standard sounds) in the violated condition reflects a violation of metrical anticipation and compared the amplitude differences between the violated and unviolated conditions within the MMN and P300 time windows. For each time window, the average amplitude differences of the two conditions were compared using a paired t-test

after confirming a normal distribution; a paired Wilcoxon test was used if normality could not be confirmed.

Finally, we explored the relationship between the neural entrainment and anticipatory processes. For neural entrainment, we focused on Bands 3 and 4, computing a band-specific spectral difference by subtracting the spectral difference in Band 4 from that in Band 3 (see **Fig. 2F**). A positive value indicates that neural activity is more strongly entrained to the triple-related feature in Band 3, whereas a negative value indicates stronger entrainment to the quadruple-related feature in the same band. We predicted that this band-specific spectral difference would reflect meter-related neural entrainment the most. For anticipatory processing, we calculated the amplitude difference between the original amplitude differences in the violated and unviolated conditions (see **Fig. 3F**), and averaged separately for the MMN and P300 time windows. Two linear models were constructed: one with the MMN amplitude difference as the response variable (see **Fig. 3G**), and the other with the P300 amplitude difference as the response variable, using the band-specific spectral difference as the explanatory variable (see **Fig. 3H**). All statistical tests were two-sided and performed using the R software (R Core Team, 2024) (version 4.1.1), with the significance level set at $\alpha = 0.05$.

Results

The present study aimed to elucidate the neural mechanisms underlying meter perception using dual-meter stimuli (**Fig. 1A,B**), which consists of band-limited noise bursts varying in center frequency and duration. This design allows the same physical sound sequence to elicit different metrical interpretations, depending on whether listeners attend to frequency or duration differences (**Fig. 1C**), thereby dissociating the effects of acoustic variability from neural processing. We compared the spectral profiles of EEG signals when participants focused on triple-meter versus quadruple-meter features to assess the neural oscillations associated with the metrical structure. We then calculated ERPs in response to disruptions of the anticipated metrical structure, implemented by altering the frequency or duration of one sound later in the stimulus. We also examined whether neural entrainment predicted the responses to metrical violations.

Neural entrainment related to meter

To examine the neural entrainment related to the meter, we first selected electrodes for analysis. We computed the EEG spectrum using data from the initial 18 noise bursts, irrespective of whether

participants attended to the center frequency or duration differences (**Fig. 2A**). Additionally, we generated a surrogate signal and calculated its spectrum. We then focused on the frequency of approximately 3.33 Hz (ranging from 2.83 to 3.83 Hz), which reflects neural synchronization to each noise burst (beat), and compared the ratio of the original to surrogate signals. The highest ratio, averaged across participants, was observed at electrode T7 (-0.02 ± 0.01; mean ± SD) (**Fig. 2B**). Based on these findings, T7 and its neighboring electrodes, FC5 (-0.03 ± 0.01), C3 (-0.03 ± 0.02), and CP5 (-0.03 ± 0.02), were selected for further analysis.

Within the same time window, we computed the EEG spectral amplitude while categorizing the data according to whether the participants attended triple- or quadruple-related acoustic features. We averaged the spectral amplitudes from T7, FC5, C3, and CP5 and defined five frequency bands corresponding to meters (**Fig. 2C**): Band-1 (beat frequency) at 3.33 Hz (ranged 3.26–3.43 Hz); Band-2 (duple meter) at 1.67 Hz (1.60–1.76 Hz); Band-3 (triple) at 1.11 Hz (1.04–1.20 Hz); Band-4 (quadruple) at 0.83 Hz (0.76–0.93 Hz); and Band-5 at 0.67 Hz (0.60–0.76 Hz). To assess whether the participants' EEG signals synchronized with the triple- or quadruple-related acoustic features of the dual-meter stimuli, we calculated the amplitude difference by subtracting the spectrum associated with attention to quadruple features from that associated with attention to triple features (**Fig. 2D**). The overall spectral difference values (mean ± SD) of Band-1 to Band-5 were $-1.14 \times 10^{-3} \pm 3.25 \times 10^{-3}$ μV, $0.14 \times 10^{-3} \pm 3.16 \times 10^{-3}$ μV, $2.30 \times 10^{-3} \pm 5.41 \times 10^{-3}$ μV, $0.39 \times 10^{-3} \pm 4.59 \times 10^{-3}$ μV, and $-0.19 \times 10^{-3} \pm 3.76 \times 10^{-3}$ μV, respectively (**Fig. 2E**). To determine whether the spectral difference varied across the frequency bands, we fitted the data to a linear mixed-effects regression model with spectral difference as the objective variable, the meter-related frequency band as a fixed effect, and participant ID as a random effect. The model revealed a significant effect of the spectral bands ($F(4, 132) = 3.52$, $p = 0.009$). Band 1 showed significantly lower spectral difference values compared to Band 3 (*estimate* $= -3.44 \times 10^{-3}$, $SE = 0.95 \times 10^{-3}$, $t(132) = -3.62$, $p_{Tukey} = 0.004$, **Table 1**). Other pairwise comparisons did not reach significance ($p_{Tukey} > 0.05$), although the contrast between Bands 3 and 5 approached significance ($p_{Tukey} = 0.073$). These results suggest that the overall spectral difference reflects neural synchronization to acoustic features that form meter perception.

MMN and P300 components related to metrical anticipation

To measure the extent of the participants' metrical anticipation, we modified the center frequency or

duration of the 19th, 20th, or 21st band-noise bursts, referred to as altered sounds, and defined standard sounds as those in which both the center frequency and duration matched the altered sounds (**Fig. 1C**). Thus, when participants focused on the altered acoustic feature, the metrical structure was violated (violated condition). In contrast, when the participants focused on the other feature (unviolated condition), the metrical structure remained intact despite the modifications.

To identify the electrodes for analyzing the neural processes associated with meter anticipation, we first calculated the ERPs for standard and altered sounds (**Fig. 3A**), and computed the amplitude difference (altered minus standard). A clear negative deflection was observed in the altered sounds relative to the standard sounds between 190 and 350 ms, which we designated as the MMN time window. We then computed the average potential difference across participants for each electrode (**Fig. 3B**). The most negative value was detected at electrode Fz (−0.90 ± 0.46 μV; mean ± SD). Consequently, Fz and its neighboring electrodes, F3 (−0.76 ± 0.36 μV), F4 (−0.90 ± 0.38 μV), FC1 (−0.75 ± 0.45 μV), FC2 (−0.89 ± 0.46 μV), and Cz (−0.75 ± 0.43 μV), were chosen for further analysis. We defined the 351–600 ms window as the P300 time window and analyzed the same electrodes.

Subsequently, we investigated the neural processes underlying anticipatory mechanisms in meter perception by calculating the ERPs for standard and altered sounds under both violated and unviolated conditions (**Fig. 3C,D**). For each condition, we computed the amplitude difference (altered minus standard) (**Fig. 3E**). For the MMN, the amplitude difference was −1.02 ± 1.51 μV in the unviolated condition and −0.65 ± 1.37 μV in the violated condition. For the P300, the potential difference was −1.03 ± 1.34 μV in the unviolated condition and −0.75 ± 1.39 μV in the violated condition. Although the P300 component appeared larger in the violated condition, no statistically significant differences were found for either MMN ($V = 351$, $p = 0.370$) or P300 ($t = 0.90$, $df = 33$, $p = 0.374$).

Predicting the anticipatory process from neural entrainment

We examined whether neural entrainment predicts anticipatory mechanisms during meter perception by focusing on the band-specific spectral difference (Band 3 minus Band 4), which we considered to best reflect neural entrainment related to meter (**Fig. 2F**). For visualization purposes, we categorized the participants into three groups based on their band-specific spectral differences in descending order: good ($n = 11$), middle ($n = 12$), and poor switchers ($n = 11$). For each group, we calculated the amplitude difference by subtracting the amplitude difference in the unviolated condition from that in the violated condition. In the MMN time window, no prominent differences were observed between

groups (**Fig. 3F**). In contrast, in the P300 time window, the good switchers exhibited larger amplitudes, followed by the middle and poor switchers (**Fig. 3F**). Linear regression models revealed that the band-specific spectral difference did not significantly predict the difference in the MMN amplitude (*slope* = 31.28, $p = 0.445$; **Fig. 3G**, **Table 2**), whereas it significantly predicted the difference in P300 amplitude (*slope* = 105.04, $p = 0.013$; **Fig. 3H**, **Table 3**). These findings suggest that attention shapes anticipation for metrical structures, enhances neural entrainment, and ultimately facilitates meter perception.

Discussion

This study aimed to clarify the neural mechanisms underlying meter perception by using dual-meter sound stimuli designed to dissociate acoustic influences from neural responses (**Fig. 1A,B**). First, we examined oscillatory neural entrainment by comparing EEG spectral profiles when participants selectively attended to either triple-related or quadruple-related features of the stimuli (**Fig. 2D**). We found that the overall spectral differences varied significantly across meter-related frequency bands, suggesting that this measure reflects meter-related neural entrainment (**Fig. 2E**). However, the specific spectral difference between the frequency bands corresponding to the triple and quadruple meters, the primary contrast of interest, was not statistically significant (**Fig. 2E,F**). We also analyzed the MMN and P300 components of ERPs elicited by deviations in metrical structure. Although differences in ERP amplitudes were observed between conditions in which metrical expectations were violated and unviolated (**Fig. 3E**), these differences were not statistically significant. We then investigated whether the degree of neural entrainment predicted the magnitude of the anticipatory responses. Our results showed that the entrainment index significantly predicted the P300 amplitude but not the MMN amplitude (**Fig. 3F–H**). These findings suggest that directing attention to specific acoustic features enhances the anticipation of metrical structures, which in turn engages neural entrainment mechanisms involved in meter perception. Previous studies have proposed that neural entrainment and anticipatory processing are two core mechanisms supporting meter perception (Harding et al., 2025; Snyder et al., 2024; Vuust et al., 2022) and that entrainment may serve as a basis for anticipation, as described by the dynamic attending theory (Ellis & Jones, 2009; Jones, 2018; Jones & Boltz, 1989). Our study advances this framework by directly demonstrating a functional relationship between neural entrainment and anticipatory processing, which was made possible by the combined use of dual-meter stimuli and EEG measurements.

The lack of significant effects in the statistical tests for the band-specific spectral differences (**Fig. 2F**) and for the amplitude differences between the conditions, indicative of the anticipatory process (**Fig. 3E**) is likely due to individual variability in top-down attentional processes. In our previous study (Kondoh et al., 2021), we observed individual differences in how participants switched their attention between the frequency and duration of dual-meter stimuli; those with prior training to tap along each noise burst while focusing on a single acoustic feature exhibited greater switching. In the present study, while the participants received pre-training to familiarize themselves with the stimuli, the main task required them to count the number of sounds associated with a given acoustic feature. Unlike the previous study, they were not explicitly instructed to attend to frequency or duration contrasts, which may have allowed individual differences in attention control to influence the results. Moreover, the use of a broader range of stimuli compared with our previous study may have made it challenging for some listeners to shift their attention between acoustic features. Additionally, musical experience may affect top-down processing of meter perception (Drake et al., 2000; Geiser et al., 2009; Kondoh et al., 2021). Because musical experience was not the primary focus of this study, we did not assess it using measures such as the Goldsmiths Musical Sophistication Index (Müllensiefen et al., 2014; Sadakata et al., 2022). Future research should investigate these factors to better explain the individual differences observed in the neural entrainment and anticipatory responses.

The band-specific spectral differences reflecting neural entrainment did not predict amplitude differences between violated and unviolated conditions in the MMN time window (**Fig. 3G**); however, they significantly predicted the amplitude differences in the P300 time window (**Fig. 3H**). This dissociation likely reflects functional distinctions between the MMN and P300 components. MMN is typically associated with relatively pre-attentive detection of deviations in acoustic features (Koelsch et al., 2019; Näätänen et al., 1978, 2007; Vuust et al., 2022). Previous studies investigating bottom-up processing in meter perception have shown that disruptions to amplitude regularity in sequences establishing a duple meter, for example, inserting a pattern like "loud-soft-loud-soft-loud-loud-soft…", can elicit MMN responses (Radchenko et al., 2023; Zhao et al., 2017). In the present study, the MMN may have primarily reflected the immediate acoustic change of a single manipulated noise burst rather than a violation of higher-order metrical expectations. By contrast, the P300 component is known to reflect conscious evaluation processes following the detection of anticipation violations (Koelsch et al., 2019; Polich, 2007). Supporting this interpretation, prior studies have shown that violations of rhythmic expectations evoke stronger P300 responses in jazz musicians than in non-musicians,

suggesting that the P300 is shaped by higher-order metrical evaluations and musical expertise (Vuust et al., 2009; Vuust & Witek, 2014). Thus, the P300 observed in this study likely reflects the violation of abstract metrical expectations rather than a response to simple acoustic deviance. This functional distinction may explain why neural entrainment predicted the P300 component but not the MMN.

Previous studies on neural entrainment related to meter perception have often relied on the amplitude of EEG signals at frequencies corresponding to the perceived meter, known as SSEPs (Chemin et al., 2014; Cirelli et al., 2016; Gibbings et al., 2023; Nave et al., 2022; Nozaradan et al., 2011, 2012, 2016, 2018). Typically, SSEP analysis involves applying FFT and calculating the amplitude at a target frequency by subtracting the average amplitude of its neighboring frequency bins. These neighboring bins serve as a baseline for isolating frequency-specific activity. However, this method can limit the accuracy of comparisons across the frequency bins. In contrast, the present study introduced a novel approach by comparing the EEG spectra between conditions in which participants attended to either the center frequency or duration of the dual-meter stimuli. This approach enables direct comparisons across meter-related frequency bands, while minimizing the confounding effects of acoustic variability. As a result, the overall spectral differences reflected meter-related neural entrainment (**Fig. 2E**). Thus, in addition to traditional SSEP techniques, our method offers an alternative perspective that may enhance our understanding of attention-modulated neural entrainment in meter perception.

This study had several limitations. First, the present study used dual-meter stimuli with IOIs fixed at 300 ms (200 beats per minute or bpm) to maintain consistency with our previous study (Kondoh et al., 2021). This setting is unlikely to be problematic for beat perception because earlier research has indicated an optimal tempo for beat perception ranging from 100 ms to 2 s in the IOI (London, 2012; Repp & Su, 2013). However, 200 bpm could be considered very fast in Western musical standards; thus, further research employing slower tempos is necessary to clarify its effects on neural entrainment and anticipatory responses during meter perception. Second, we did not rigorously assess the relationship between band-specific spectral differences (reflecting neural synchronization related to meters), amplitude differences between conditions (reflecting expectancy violations), and behavioral measures of meter perception. As an exploratory analysis, we examined the correlations between task sensitivity, measured by the ability to detect whether an altered sound disrupted metrical structures, and EEG indicators (**Text S1**, **Fig. S3**). However, because we did not have a priori hypotheses about specific electrode sites, we only reported electrodes with high correlation coefficients. Future studies

should test the relationship between EEG and behavioral indicators more systematically, using the electrodes identified in our supplementary analysis.

In conclusion, this study investigated the neural basis of meter perception by using dual-meter stimuli designed to minimize the confounding effects of acoustic variation on neural activity. We hypothesized that top-down selective attention to either the frequency or duration of a sound sequence would modulate the bottom-up processing of that feature, leading to neural synchronization with acoustic regularities and the formation of corresponding expectations that support meter perception. Consistent with this hypothesis, we found that the strength of meter-related neural entrainment predicted the strength of anticipation. These findings represent a methodological advance in isolating neural responses to meters from acoustic confounds, and provide direct evidence linking neural synchronization to anticipatory processing in meter perception. We believe that these results mark an important step toward a deeper understanding of meter perception, a fundamental phenomenon in music.

Conflict of interest

The authors have declared that no competing interests exist.

Acknowledgments

This work was supported by grants from JSPS KAKENHI (grant numbers 23H05428 to KO, 23K18475 and 24H00735 to ROT, and 24KJ1930 to SK). The authors would like to thank Drs. Hironori Nakatani and Reiko Shiba for their valuable feedback on the early versions of our pilot experiment. We would also like to thank Dr. Shinya Fujii and the members of his laboratory for their insightful comments on the preliminary results of this study.

Author Contributions

Conceptualization: SK, KO, ROT

Data Curation: SK

Formal Analysis: SK, ROT

Funding Acquisition: SK, KO, ROT

Investigation: SK

Methodology: SK, ROT

Project Administration: KO, ROT

Software: SK, ROT

Supervision: KO, ROT

Visualization: SK, ROT

Writing – Original Draft Preparation: SK

Writing – Review & Editing: SK, KO, ROT

Data Availability

The data supporting the findings of this study as well as the sound files of the dual-meter stimuli used are available from the Open Science Framework repository (https://osf.io/hw2y8).

Reference

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

Bregman, A. S. (1994). *Auditory scene analysis: The perceptual organization of sound*. MIT press.

Bregman, A. S., Ahad, P. A., & Van Loon, C. (2001). Stream segregation of narrow-band noise bursts. *Perception and Psychophysics*, *63*(5), 790–797. https://doi.org/10.3758/BF03194438

Chemin, B., Mouraux, A., & Nozaradan, S. (2014). Body movement selectively shapes the neural representation of musical rhythms. *Psychological Science*, *25*(12), 2147–2159. https://doi.org/10.1177/0956797614551161

Chen, J. L., Penhune, V. B., & Zatorre, R. J. (2009). The role of auditory and premotor cortex in sensorimotor transformations. *Annals of the New York Academy of Sciences*, *1169*, 15–34. https://doi.org/10.1111/j.1749-6632.2009.04556.x

Chen, J. L., Zatorre, R. J., & Penhune, V. B. (2006). Interactions between auditory and dorsal premotor cortex during synchronization to musical rhythms. *NeuroImage*, *32*(4), 1771–1781. https://doi.org/10.1016/j.neuroimage.2006.04.207

Cherry, E. C. (1953). Some experiments on the recognition of speech, with one and with two ears. *Journal of the Acoustical Society of America*, *25*, 975–979. https://doi.org/10.1121/1.1907229

Cirelli, L. K., Spinelli, C., Nozaradan, S., & Trainor, L. J. (2016). Measuring Neural Entrainment to Beat and Meter in Infants: Effects of Music Background. *Frontiers in Neuroscience*, *10*, 229. https://doi.org/10.3389/fnins.2016.00229

Cooper, G., & Meyer, L. B. (1963). *The rhythmic structure of music*. University of Chicago press.

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. https://doi.org/10.1016/j.jneumeth.2003.10.009

Delorme, A., Sejnowski, T., & Makeig, S. (2007). Improved rejection of artifacts from EEG data using high-order statistics and independent component analysis. *NeuroImage*, *34*, 1443–1449.

Drake, C., Penel, A., & Bigand, E. (2000). Tapping in time with mechanically and expressively performed music. *Music Perception*, *18*(1), 1–23. https://doi.org/10.2307/40285899

Ellis, R. J., & Jones, M. R. (2009). The role of accent salience and joint accent structure in meter perception. *Journal of Experimental Psychology. Human Perception and Performance*, *35*(1), 264–280. https://doi.org/10.1037/a0013482

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/BF03193146

Fujii, S., & Schlaug, G. (2013). The Harvard Beat Assessment Test (H-BAT): a battery for assessing beat perception and production and their dissociation. *Frontiers in Human Neuroscience*, *7*, 771. https://doi.org/10.3389/fnhum.2013.00771

Fujioka, T., Trainor, L. J., Large, E. W., & Ross, B. (2009). Beta and gamma rhythms in human auditory cortex during musical beat processing. *Annals of the New York Academy of Sciences*, *1169*, 89–92. https://doi.org/10.1111/j.1749-6632.2009.04779.x

Geiser, E., Ziegler, E., Jancke, L., & Meyer, M. (2009). Early electrophysiological correlates of meter and rhythm processing in music perception. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*, *45*(1), 93–102. https://doi.org/10.1016/j.cortex.2007.09.010

Gibbings, A., Henry, M. J., Cruse, D., Stojanoski, B., & Grahn, J. A. (2023). Attention modulates neural measures associated with beat perception. *The European Journal of Neuroscience*, *57*(9), 1529–1545. https://doi.org/10.1111/ejn.15962

Harding, E. E., Kim, J. C., Demos, A. P., Roman, I. R., Tichko, P., Palmer, C., & Large, E. W. (2025). Musical neurodynamics. *Nature Reviews. Neuroscience*, *26*, 293–307.

https://doi.org/10.1038/s41583-025-00915-4

Henninger, F., Shevchenko, Y., Mertens, U. K., Kieslich, P. J., & Hilbig, B. E. (2022). lab.js: A free, open, online study builder. *Behavior Research Methods*, *54*(2), 556–573. https://doi.org/10.3758/s13428-019-01283-5

Honing, H. (2013). Structure and interpretation of rhythm in music. In D. Deutsch (Ed.), *The psychology of music* (3rd ed., pp. 369–404). Academic Press. https://doi.org/10.1016/B978-0-12-381460-9.00009-2

Jones, M. R. (2018). *Time will tell: A theory of dynamic attending*. Oxford University Press.

Jones, M. R., & Boltz, M. (1989). Dynamic attending and responses to time. *Psychological Review*, *96*(3), 459–491. https://doi.org/10.1037/0033-295X.96.3.459

Koelsch, S., Vuust, P., & Friston, K. (2019). Predictive Processes and the Peculiar Case of Music. *Trends in Cognitive Sciences*, *23*(1), 63–77. https://doi.org/10.1016/j.tics.2018.10.006

Kondoh, S., Okanoya, K., & Tachibana, R. O. (2021). Switching perception of musical meters by listening to different acoustic cues of biphasic sound stimulus. *PloS One*, *16*(8), e0256712. https://doi.org/10.1371/journal.pone.0256712

Kotz, S. A., Ravignani, A., & Fitch, W. T. (2018). The evolution of rhythm processing. *Trends in Cognitive Sciences*, *22*(10), 896–910. https://doi.org/10.1016/j.tics.2018.08.002

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, *82*(13). https://doi.org/10.18637/jss.v082.i13

Large, E. W., & Jones, M. R. (1999). The dynamics of attending: How people track time-varying events. *Psychological Review*, *106*(1), 119–159. https://doi.org/10.1037//0033-295x.106.1.119

Lenth, R. V. (2024). *emmeans: Estimated Marginal Means, aka Least-Squares Means*. https://CRAN.R-project.org/package=emmeans

Lerdahl, F., & Jackendoff, R. S. (1983). *A generative theory of tonal music*. MIT press.

London, J. (2012). *Hearing in time: Psychological aspects of musical meter* (2nd ed.). Oxford University Press.

Makeig, S., Debener, S., Onton, J., & Delorme, A. (2004). Mining event-related brain dynamics. *Trends in Cognitive Sciences*, *8*(5), 204–210.

Müllensiefen, D., Gingras, B., Musil, J., & Stewart, L. (2014). Measuring the facets of musicality: The Goldsmiths Musical Sophistication Index (Gold-MSI). *Personality and Individual Differences*, *60*, S35. https://doi.org/10.1016/j.paid.2013.07.081

Näätänen, R., Gaillard, A. W. K., & Mäntysalo, S. (1978). Early selective-attention effect on evoked potential reinterpreted. *Acta Psychologica*, *42*(4), 313–329. https://doi.org/10.1016/0001-6918(78)90006-9

Näätänen, R., Paavilainen, P., Rinne, T., & Alho, K. (2007). The mismatch negativity (MMN) in basic research of central auditory processing: A review. *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology*, *118*(12), 2544–2590. https://doi.org/10.1016/j.clinph.2007.04.026

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. https://doi.org/10.1111/psyp.13963

Nozaradan, S., Keller, P. E., Rossion, B., & Mouraux, A. (2018). EEG Frequency-Tagging and Input–Output Comparison in Rhythm Perception. *Brain Topography*, *31*(2), 153–160. https://doi.org/10.1007/s10548-017-0605-8

Nozaradan, S., Peretz, I., & Keller, P. E. (2016). Individual Differences in Rhythmic Cortical Entrainment Correlate with Predictive Behavior in Sensorimotor Synchronization. *Scientific Reports*, *6*, 20612. https://doi.org/10.1038/srep20612

Nozaradan, S., Peretz, I., Missal, M., & Mouraux, A. (2011). Tagging the neuronal entrainment to beat and meter. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *31*(28), 10234–10240. https://doi.org/10.1523/JNEUROSCI.0411-11.2011

Nozaradan, S., Peretz, I., & Mouraux, A. (2012). Selective neuronal entrainment to the beat and meter embedded in a musical rhythm. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *32*(49), 17572–17581. https://doi.org/10.1523/JNEUROSCI.3203-12.2012

Palmer, C., & Krumhansl, C. L. (1990). Mental representations for musical meter. *Journal of*

Experimental Psychology. Human Perception and Performance, *16*(4), 728–741. https://doi.org/10.1037/0096-1523.16.4.728

Patel, A. D. (2006). Musical rhythm, linguistic rhythm, and human evolution. *Music Perception*, *24*(1), 99–104. https://doi.org/10.1525/mp.2006.24.1.99

Patel, A. D. (2010). *Music, language, and the brain*. Oxford university press.

Pion-Tonachini, L., Kreutz-Delgado, K., & Makeig, S. (2019). ICLabel: An automated electroencephalographic independent component classifier, dataset, and website. *NeuroImage*, *198*, 181–197. https://doi.org/10.1016/j.neuroimage.2019.05.026

Polich, J. (2007). Updating P300: an integrative theory of P3a and P3b. *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology*, *118*(10), 2128–2148. https://doi.org/10.1016/j.clinph.2007.04.019

R Core Team. (2024). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Radchenko, G., Demareva, V., Gromov, K., Zayceva, I., Rulev, A., Zhukova, M., & Demarev, A. (2023). Neural mechanisms of temporal and rhythmic structure processing in non-musicians. *Frontiers in Neuroscience*, *17*, 1124038. https://doi.org/10.3389/fnins.2023.1124038

Repp, B. H., & Su, Y. H. (2013). Sensorimotor synchronization: A review of recent research (2006-2012). *Psychonomic Bulletin & Review*, *20*(3), 403–452. https://doi.org/10.3758/s13423-012-0371-2

Sadakata, M., Yamaguchi, Y., Ohsawa, C., Matsubara, M., Terasawa, H., von Schnehen, A., Müllensiefen, D., & Sekiyama, K. (2022). The Japanese translation of the Gold-MSI: Adaptation and validation of the self-report questionnaire of musical sophistication. *Musicae Scientiae: The Journal of the European Society for the Cognitive Sciences of Music*. https://doi.org/10.1177/10298649221110089

Snyder, J. S., Gordon, R. L., & Hannon, E. E. (2024). Theoretical and empirical advances in understanding musical rhythm, beat and metre. *Nature Reviews Psychology*, 1–14. https://doi.org/10.1038/s44159-024-00315-y

Snyder, J. S., & Large, E. W. (2005). Gamma-band activity reflects the metric structure of rhythmic tone sequences. *Cognitive Brain Research*, *24*(1), 117–126.

https://doi.org/10.1016/j.cogbrainres.2004.12.014

Sutton, S., Braren, M., Zubin, J., & John, E. R. (1965). Evoked-Potential Correlates of Stimulus Uncertainty. *Science*, *150*(3700), 1187–1188. https://doi.org/10.1126/science.150.3700.1187

Vuust, P., Heggli, O. A., Friston, K. J., & Kringelbach, M. L. (2022). Music in the brain. *Nature Reviews. Neuroscience*, *23*(5), 287–305. https://doi.org/10.1038/s41583-022-00578-5

Vuust, P., Ostergaard, L., Pallesen, K. J., Bailey, C., & Roepstorff, A. (2009). Predictive coding of music - Brain responses to rhythmic incongruity. *Cortex*, *45*(1), 80–92. https://doi.org/10.1016/j.cortex.2008.05.014

Vuust, P., & Witek, M. A. G. (2014). Rhythmic complexity and predictive coding: A novel approach to modeling rhythm and meter perception in music. *Frontiers in Psychology*, *5*, 1111. https://doi.org/10.3389/fpsyg.2014.01111

Woods, K. J. P., Siegel, M. H., Traer, J., & McDermott, J. H. (2017). Headphone screening to facilitate web-based auditory experiments. *Attention, Perception & Psychophysics*, *79*(7), 2064–2072. https://doi.org/10.3758/s13414-017-1361-2

Zatorre, R. J., Chen, J. L., & Penhune, V. B. (2007). When the brain plays music: auditory–motor interactions in music perception and production. *Nature Reviews. Neuroscience*, *8*(7), 547–558. https://doi.org/10.1038/nrn2152

Zhao, T. C., Lam, H. T. G., Sohi, H., & Kuhl, P. K. (2017). Neural processing of musical meter in musicians and non-musicians. *Neuropsychologia*, *106*, 289–297. https://doi.org/10.1016/j.neuropsychologia.2017.10.007

Figures

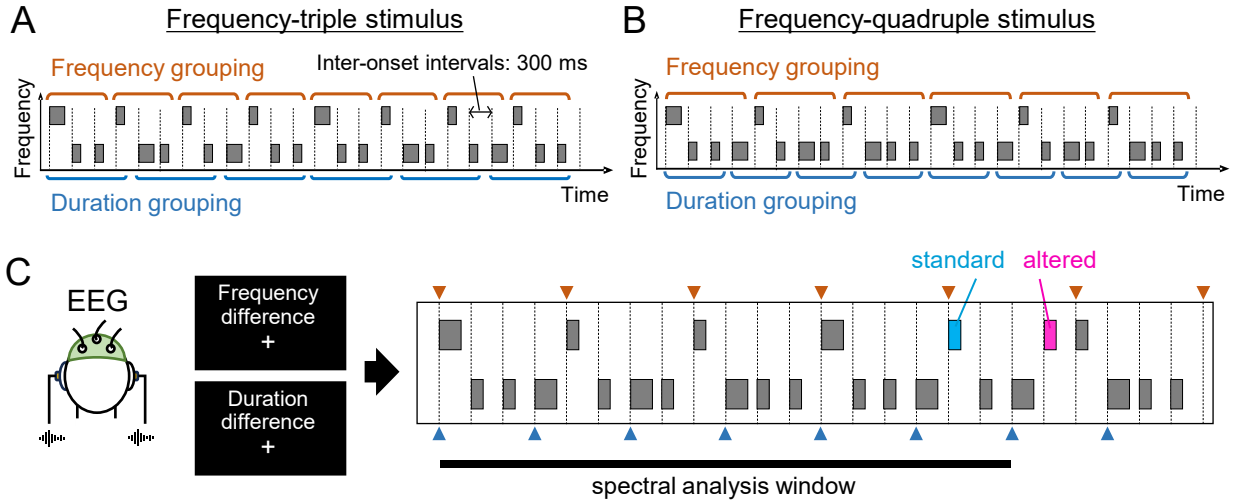

Fig 1. Schematic drawing of dual-meter stimuli and experimental design. A. Example stimulus from the frequency-triple category with "high-low-low-..." frequency and "long-short-short-short-…" duration patterns. **B.** Example stimulus from the frequency-quadruple category with "high-low-low-low..." frequency and "long-short-short-…" duration patterns. The center frequency of the higher sound was 1600 Hz and that of the lower sound was 1400 Hz, resulting in a frequency difference of 200 Hz. The duration of the longer sound was 150 ms, whereas that of the shorter sound was 50 ms, creating a duration difference of 100 ms. The bandwidth of each noise burst was 120% of its center frequency, and the inter-onset intervals (IOIs) were fixed at 300 ms. **C.** Experimental procedure. Participants attended to either the duration or frequency differences of the stimulus during electroencephalogram (EEG) recordings. Spectral analysis was performed on the EEG data corresponding to the initial 18 sounds (indicated by the black line). In this example, the 20th noise burst was altered by increasing its center frequency from low to high (altered sound is indicated by the pink square). We defined standard sounds as those that matched the altered sounds in center frequency and duration, presenting at least three sounds earlier (indicated by the light blue square). When focusing on the frequency difference, the metrical structure of the quadruple meter (indicated by orange triangles) is violated by the altered sound (violated condition). In contrast, when focusing on the duration difference, the metrical structure of the triple meter (indicated by blue triangles) remains intact (unviolated condition). Event-related potentials (ERPs) were calculated around the altered and standard sounds.

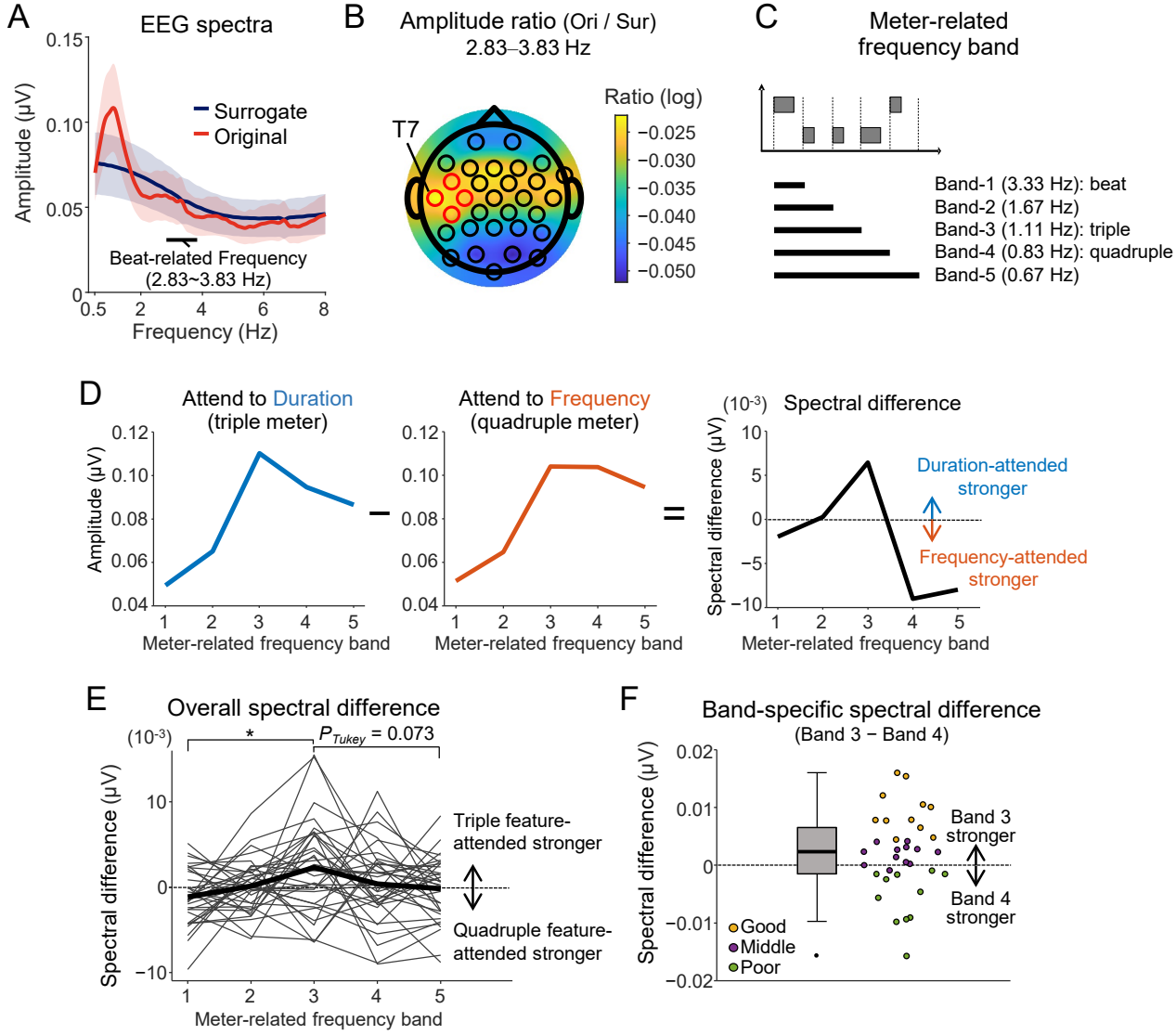

Fig. 2 Neural index reflecting oscillatory entrainment associated with meter perception. A. EEG spectra of the original and surrogate signals across all electrodes without differentiating whether participants focused on triple- or quadruple-related acoustic features. The shaded areas represent the standard deviations (SD) across the electrodes. The analysis targeted a beat-related frequency range (2.83–3.83 Hz). **B.** Topographical electrode distribution of the log-transformed ratio of the original to surrogate EEG signal amplitudes in the beat-related frequency range. The T7 electrode showed the highest ratio, and the adjacent electrodes, FC5, C3, and CP5, were selected for further analysis (indicated by red circles). **C.** Description of meter-related frequency bands: Band-1 (3.33 [3.26–3.43] Hz) corresponding to beat, Band-2 (1.67 [1.60–1.76] Hz), Band-3 (1.11 [1.04–1.20] Hz) corresponding to the triple meter, Band-4 (0.83 [0.76–0.93] Hz) corresponding to the quadruple meter, and Band-5 (0.67 [0.60–0.76] Hz). **D.** Spectral analysis to calculate neural entrainment (examples from one participant). We calculated the spectral amplitude for the triple-related (left) and quadruple-related features (middle) and computed the difference (triple minus quadruple) to eliminate the effects of acoustic variability (right). **E.** Overall spectral differences between triple-related and quadruple-related acoustic features across meter-related frequency bands. Each grey line represents individual data, whereas the bold black line indicates the average across participants. Positive values indicate stronger entrainment to triple-related features, whereas negative values indicate stronger entrainment to quadruple-related features. Asterisk (*) indicates a significant difference ($p_{Tukey} < 0.05$). **F.** Box plots and individual data of the band-specific spectral difference between Bands 3 and 4. Greater values indicate stronger entrainment to triple-related features in Band-3 and weaker entrainment to quadruple-related features in the same band. The participants were classified into three groups in descending order: good (n = 11), middle (n = 12), and poor switchers (n = 11).

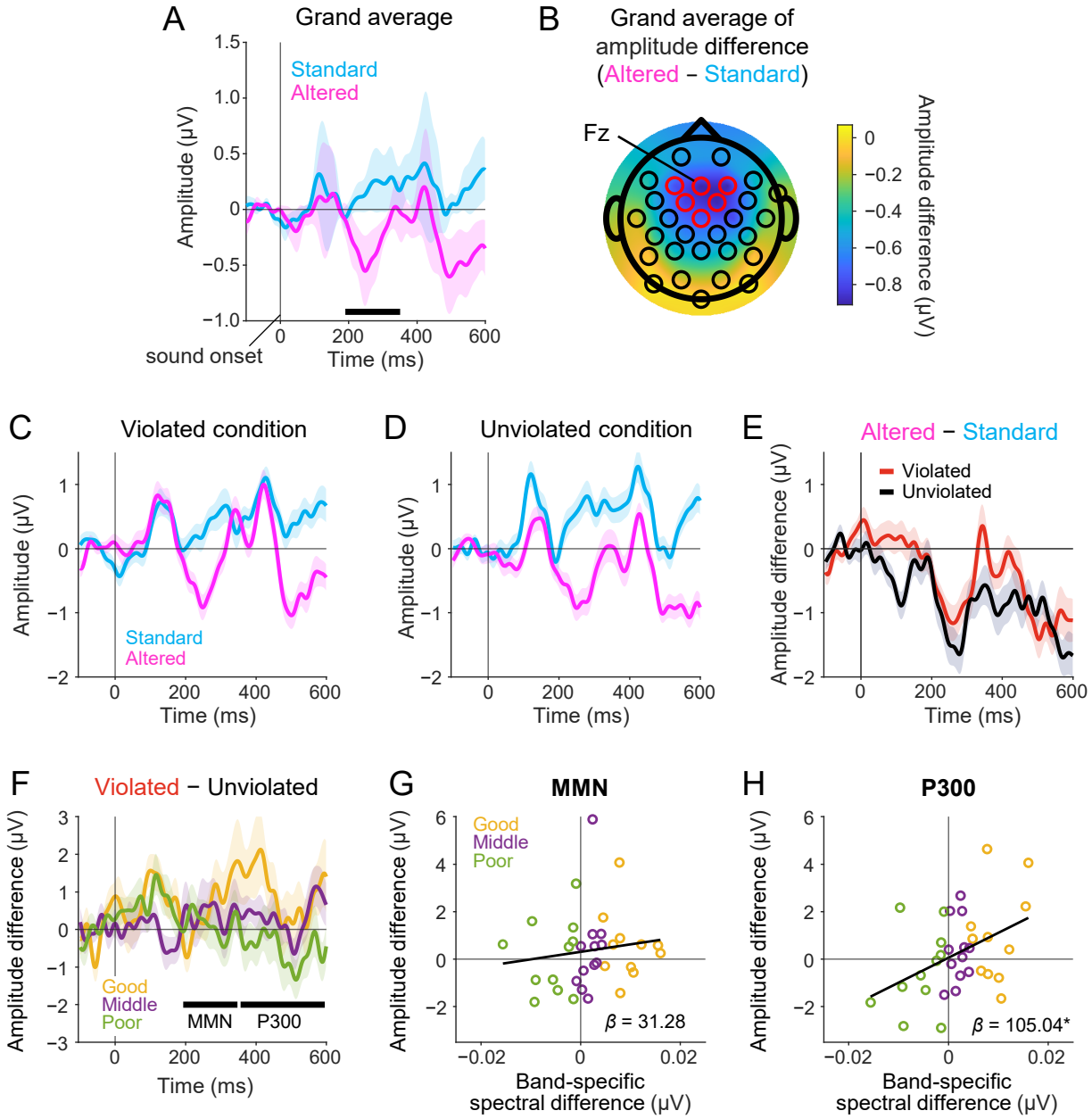

Fig. 3. Neural responses to metrical anticipation and their relationship with oscillatory entrainment. A. Grand average of ERPs for standard and altered sounds across electrodes, without categorizing violated and unviolated conditions. The shaded areas represent SD across the electrodes. The MMN time window (190–350 ms, indicated by the black line) was determined by identifying the negative deflection in the amplitude difference between altered and standard sounds. **B.** The topographical electrode map of the averaged amplitude difference (altered − standard) within 190–350 ms shows the largest negative value at the Fz electrode. Consequently, Fz and its neighboring electrodes (F3, F4, FC1, FC2, and Cz, indicated by red circles) were selected for subsequent MMN analysis. Because the P300 typically follows the MMN, the same electrodes were used for the P300 analysis, with the time window set at 351–600 ms. **C,D.** ERPs in the violated (C) and unviolated (D) conditions comparing standard and altered sounds. **E.** Amplitude differences (altered − standard sounds) under both conditions. **F.** The amplitude difference (violated − unviolated) across the three participant groups was determined in the analysis of the band-specific amplitude difference (see Fig. 2F). The shaded area in panels C–F represents the standard error across participants. **G,H.** Scatter plots showing the relationship between the band-specific amplitude difference and the amplitude difference (violated − unviolated) for the MMN (G) and P300 (H) time windows. Colored dots represent individuals in each participant group, and the β-value (slope) of the regression models is indicated. Asterisk (*) indicates a significant slope ($p < 0.05$).

Tables

Table1. Pairwise comparison of overall spectral differences between meter-related frequency bands

Contrast	*Estimate* (10^{-3})	Lower 95% CI (10^{-3})	Upper 95% CI (10^{-3})	*t*-value	p_{Tukey}
Band-1 minus Band-2	–1.29	–3.91	1.34	−1.36	0.658
Band-1 minus Band-3	–3.44	–6.07	–0.81	−3.62	0.004*
Band-1 minus Band-4	–1.53	–4.16	1.10	−1.61	0.493
Band-1 minus Band-5	–0.95	–3.58	1.67	−1.00	0.853
Band-2 minus Band-3	–2.15	–4.78	0.47	−2.27	0.162
Band-2 minus Band-4	–0.24	–2.87	2.38	−0.26	1.000
Band-2 minus Band-5	0.33	–2.29	2.96	0.35	0.997
Band-3 minus Band-4	1.91	–0.72	4.54	2.01	0.266
Band-3 minus Band-5	2.49	–0.14	5.11	2.62	0.073
Band-4 minus Band-5	0.58	–2.05	3.20	0.61	0.974

Note: CI denotes confidence interval. $SE = 0.95 \times 10^{-3}$, $df = 132$, $*p_{Tukey} < 0.05$

Table 2. Linear regression model to examine the impact of band-specific spectral differences on the amplitude difference (violated minus unviolated conditions) in the MMN time window.

Variable	*slope* [95%CI Lower, Upper]	*SE*	*t*-value	*p*-value
(Intercept)	0.31 [−0.28, 0.91]	0.29	1.08	0.290
Band-specific spectral difference	31.28 [−51.07, 113.63]	40.43	0.77	0.445

Note: CI denotes confidence interval. F (1, 32) = 0.60, Adjusted R^2= −0.01

Table 3. Linear regression model to examine the impact of band-specific spectral differences on the amplitude difference (violated minus unviolated conditions) in the P300 time window.

Variable	*slope* [95%CI Lower, Upper]	*SE*	*t*-value	*p*-value
(Intercept)	0.07 [−0.51, 0.65]	0.29	0.25	0.807
Band-specific spectral difference	105.04 [24.14, 185.94]	39.72	2.65	0.013*

Note: CI denotes confidence interval. F (1, 32) = 7.00, Adjusted R^2= 0.15, $*p < 0.05$

Supplemental Information

Text S1. Exploring the relationship between neural entrainment, anticipation, and behavioral sensitivity in detecting metrical violations.

To analyze the relationship between neural indices (entrainment and anticipation) and behavioral sensitivity to violations of metrical expectations, we conducted a supplemental behavioral analysis. After the EEG session, participants completed a task in which they judged whether an altered sound introduced a deviation when attending to either its center frequency or its duration. Each participant evaluated 28 stimuli per acoustic feature, resulting in 56 randomized trials. Participants had 2.5 s to respond and were instructed to answer as quickly and accurately as possible. If no responses were recorded, the next trial began automatically.

Sensitivity to metrical violations was quantified using the signal-detection theory (Gescheider, 1997). A hit was defined as the correct detection of a violation, whereas a miss indicated a failure to detect it. A false alarm referred to an incorrect report of a violation when none was present, and a correct rejection was an accurate identification of no violation. We define Z_N as the z-transformed value of (1 minus the proportion of false alarms), and Z_{SN} as the z-transformed value of (1 minus the proportion of hits). Detection accuracy (d') was calculated using the following equation:

$$d' = Z_N - Z_{SN} \quad \text{(Eq.S1)}$$

We then examined the relationship between d' (*mean* = 0.29, *SD* = 0.47) and three neural measures: (i) band-specific spectral differences, (ii) amplitude differences (between the violated and unviolated conditions) during the mismatch-negativity (MMN) time window, and (iii) amplitude differences in the P300 time window. Because we had no predefined hypotheses about electrode locations, we computed the Pearson correlation coefficients without formal significance testing.

Stronger positive correlations between d' and band-specific spectral differences were observed at F7 ($r = 0.38$), FC6 ($r = 0.32$), and FC5 ($r = 0.29$) (**Fig. S3A**). During the MMN window, the left hemisphere electrodes showed relatively strong negative correlations between d' and the amplitude difference at CP5 ($r = -0.30$), P3 ($r = -0.29$), and FC1 ($r = -0.25$) (**Fig. S3B**). In the P300 window, stronger positive correlations were observed in the right hemisphere electrodes such as FT10 ($r = 0.40$), PO10 ($r = 0.35$), and T8 ($r = 0.34$) (**Fig. S3C**). Therefore, these electrode sites may serve as promising targets for future studies investigating how neural indices of entrainment and anticipation relate to behavioral indices of meter violation detection.

References

Gescheider, G. A. (1997). *Psychophysics: The Fundamentals* (3rd ed.). Psychology Press.

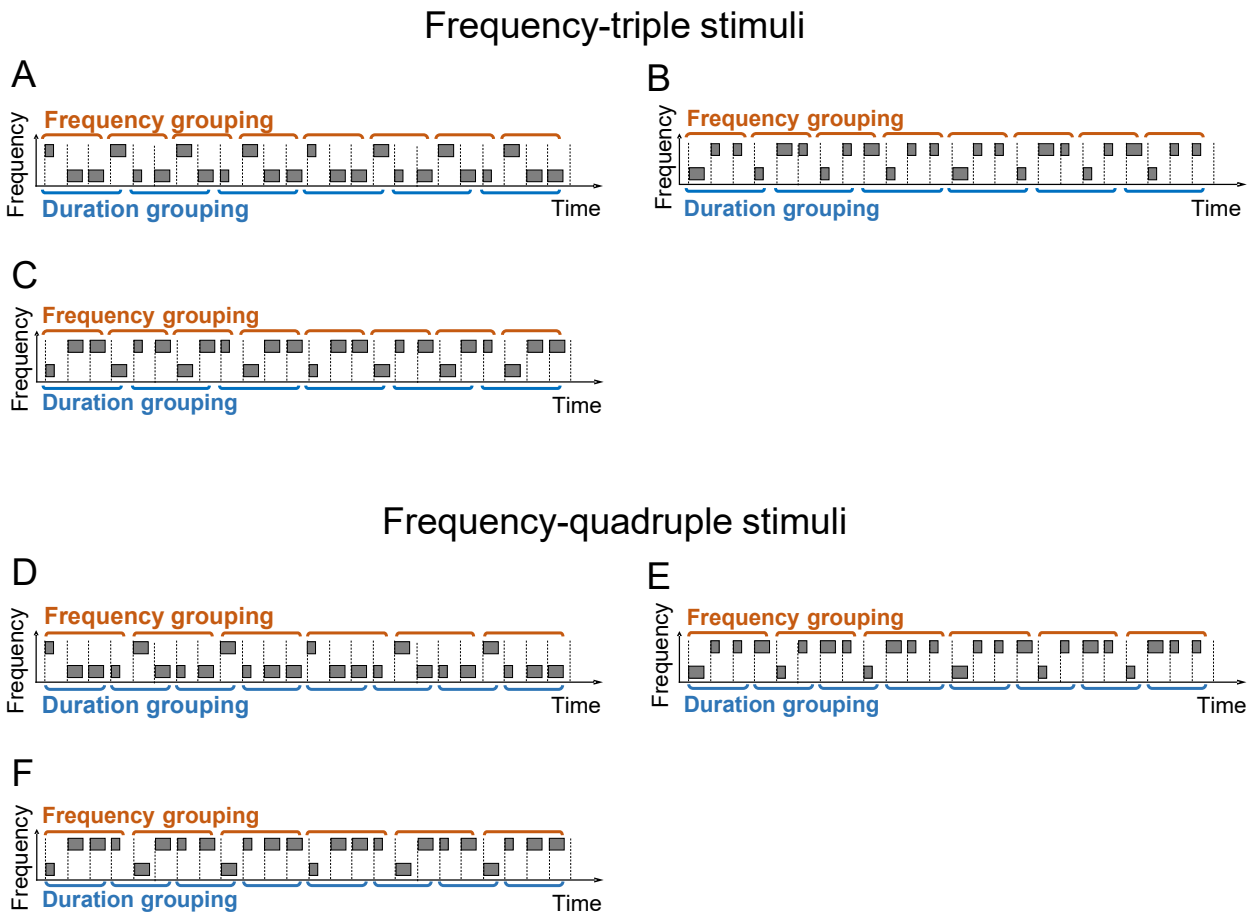

Fig. S1 Schematic representation of biphasic stimuli not included in the main text. Each gray square represents a band-limited noise burst, with each stimulus consisting of 24 bursts. The center frequency of the higher sound was 1600 Hz and that of the lower sound was 1400 Hz, resulting in a frequency difference of 200 Hz. The duration of the longer sound was 150 ms, whereas that of the shorter sound was 50 ms, creating a duration difference of 100 ms. The bandwidth of each noise burst was set to 120% of its center frequency, and the inter-onset intervals (IOIs) were fixed at 300 ms. **A-C.** Frequency-triple stimuli. The frequency pattern is either “high–low–low–...” (panel A) or “low–high–high–...” (panels B and C), while the duration pattern follows either “short–long–long–long–...” (panels A and C) or “long–short–short–short–...” (panel B). **D-F.** Frequency-quadruple stimuli. The frequency pattern follows “high–low–low–low–...” (panel D) or “low–high–high–high–...” (panels E and F), while the duration pattern is either “short–long–long–...” (panels D and F) or “long–short–short–...” (panel E).

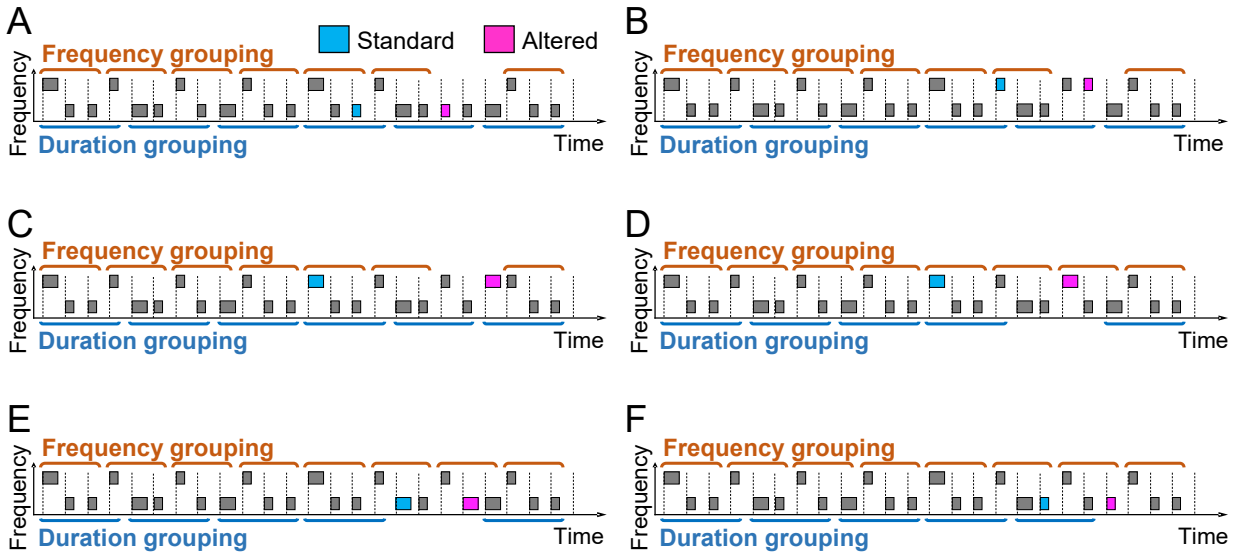

Fig. S2 Variations in one stimulus (Fig. 1A in the main text). Each pink square represents an altered sound, either by changing the center frequency of the 19th (panel A: high to low), 20th (panel B: low to high), or 21st (panel C: low to high) noise bursts or by changing the duration of the 19th (panel D: short to long), 20th (panel E: short to long), or 21st (panel F: long to short) noise bursts. For EEG analysis, standard sounds were defined as those that matched the altered sounds in both center frequency and duration and were presented at least three sounds ahead (indicated by light blue squares).

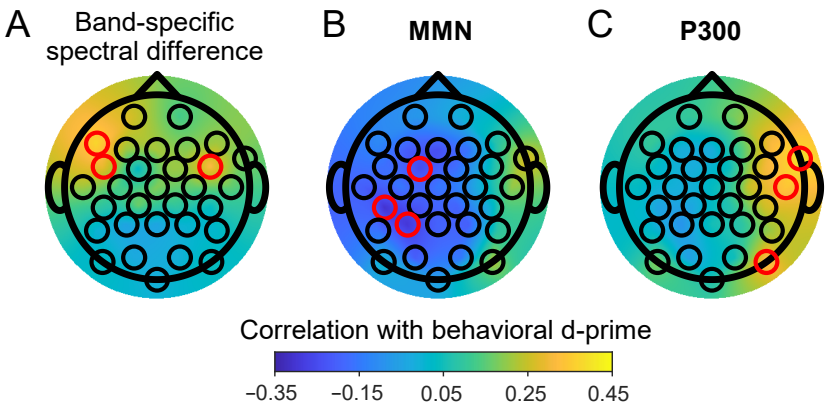

Fig. S3 Topographical distribution of correlations between behavioral d-prime (d') and EEG indices across the electrodes. A. Band-specific spectral difference. The three electrodes (F7, FC5, and FC6, indicated by red circles) showed stronger positive correlations. **B.** Amplitude difference in the MMN time window. The three electrodes (CP5, P3, and FC1, indicated by red circles) exhibited stronger negative correlations. **C.** Amplitude difference in the P300 time window. The three electrodes (FT10, PO10, and T8, indicated by red circles) showed stronger positive correlations.