Received: 18 March 2021 | Revised: 7 February 2022 | Accepted: 8 February 2022

DOI: 10.1111/ejn.15630

SPECIAL ISSUE ARTICLE

EJN European Journal of Neuroscience FENS WILEY

Neural entrainment via perceptual inferences

Alessandro Tavano[1] | **Burkhard Maess**[2] | **David Poeppel**[3,4] | **Erich Schröger**[5]

[1]Department of Neuroscience, Max Planck Institute for Empirical Aesthetics, Frankfurt am Main, Germany

[2]MEG Unit, Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany

[3]Department of Psychology, New York University, New York, New York, USA

[4]Ernst Strüngmann Institute for Neuroscience, Frankfurt am Main, Germany

[5]Wilhelm Wundt Institute for Psychology, Leipzig University, Leipzig, Germany

Correspondence
Alessandro Tavano, Department of Neuroscience, Max Planck Institute for Empirical Aesthetics, Grüneburgweg 14, 60322 Frankfurt am Main, Germany.
Email: alessandro.tavano@ae.mpg.de; alessandro.tavano.office@gmail.com

Funding information
Deutsche Forschungsgemeinschaft

Edited by: Christian Keitel

Abstract

Entrainment depends on sequential neural phase reset by regular stimulus onset, a temporal parameter. Entraining to sequences of identical stimuli also entails stimulus feature predictability, but this component is not readily separable from temporal regularity. To test if spectral regularities concur with temporal regularities in determining the strength of auditory entrainment, we devised sound sequences that varied in conditional perceptual inferences based on deviant sound repetition probability: strong inference (100% repetition probability: If a deviant appears, then it will repeat), weak inference (75% repetition probability) and no inference (50%: A deviant may or may not repeat with equal probability). We recorded EEG data from 15 young human participants pre-attentively listening to the experimental sound sequences delivered either isochronously or anisochronously (±20% jitter), at both delta (1.67 Hz) and theta (6.67 Hz) stimulation rates. Strong perceptual inferences significantly enhanced entrainment at either stimulation rate and determined positive correlations between precision in phase distribution at the onset of deviant trials and entrained power. We conclude that both spectral predictability and temporal regularity govern entrainment via neural phase control.

KEYWORDS
conditional inference, entrainment, neural phase, prediction, spectrotemporal regularities, temporal expectations

1 | INTRODUCTION

The alert human brain is often described as 'desynchronized' to distinguish its awake activity patterns from the large synchronous states occurring during sleep (Ahmed & Cash, 2013). Yet the desynchronized brain spontaneously generates rhythmic activity across different characteristic frequencies (Keitel & Gross, 2016). Furthermore, when external periodic stimulation—a flashing light, a sound sequence—is used to 'hijack' the alert brain's rhythmic mode (Keitel et al., 2014), brain activity becomes phase aligned to the rate of regular stimulus presentation. This phenomenon, termed neural entrainment, is assumed to depend on the precision of the temporal parameter in input, as neural phase aligns with the stimulus stream (Lakatos et al., 2019). Phase alignment is preeminently controlled by periodic phase reset caused by successive stimulus onset at regular times, and because the phase of neuronal

Abbreviations: EEG, electroencephalogram; FDR, false discovery rate; FFT, fast Fourier transform.

oscillations closely reflects the underlying rhythmic fluctuation of neuronal excitability, entrainment determines a gain in stimulus processing relative to aperiodic stimulus sequences (Lakatos et al., 2005). Entrained neural oscillations also make it computationally efficient to predict the onset of the next stimulus. Hence, entrainment has been proposed as a mechanism for the cyclic deployment of attention in time, aligning the excitability of neuronal ensembles to stimulus onset probability (Lakatos et al., 2008; Obleser & Kayser, 2019).

However, the repetition of identical stimuli entails stimulus feature predictability, raising the issue of whether temporal regularity may be a necessary, but not sufficient component to account for variance in neural entrainment. Lakatos et al. (2013) and O'Connell et al. (2014) working on animal data first showed that attention to external rhythmic sound sequences results in two main spectral effects: (1) amplification of neural responses in A1 to the stimulation frequency (frequency tuning) at predicted sound onset times and (2) concurrent suppression of activity for ignored frequencies. Hence, the authors suggested that entrainment should be best understood as a neural spectrotemporal sensory filter, rather than just as a temporal filter. Evidence that auditory prediction is spectrotemporal in nature also comes from the frequency tagging approach for the study of cognition (Keitel et al., 2014), according to which periodic stimulation activates stimulus-specific neuronal ensembles (Rossion, 2014). Furthermore, recent evidence in nonhuman primates in a pre-attentive setting suggests that entrainment in the delta band can be internally (i.e., top-down) guided by the repetition (grouping) of random acoustic patterns, in the absence of any cues as to pattern onset/offset (Barczak et al., 2018).

We aimed at contributing to this line of investigation by testing the sensitivity of neural entrainment to conditional perceptual inferences in audition. We parametrized spectral predictability as a conditional inference rule based on deviant sound repetition and assumed that under high deviant repetition probability entrainment to regular sound onset would be less perturbed by deviancy onset, that is, spectral predictability should reduce uncertainty in encoding the temporal component of entrainment. We asked young human participants to pre-attentively listen to oddball sound sequences carrying block-wise evidence for varying degrees of deviant sound repetition probability: *strong inferential rule*, with 100% repetition probability ('if a deviant appears, then it will repeat'); *weak inferential rule*, with 75% probability of deviant repetition; and *no inferential rule*, with 50% probability of deviant repetition. The onset of the first deviants of a pair—and, if licenced by the condition, of single deviants—within each sound sequence was made unpredictable by pseudorandomly determining the number of standards preceding first deviants (minimal interval of two standards). Implicit statistical learning was the driving force driving perceptual inferencing in audition, as participants' attention was captured by a silenced video (see Figure 1).

If entrainment relies on both temporal regularity and stimulus predictability, then an interaction should emerge favouring isochronous sequences carrying

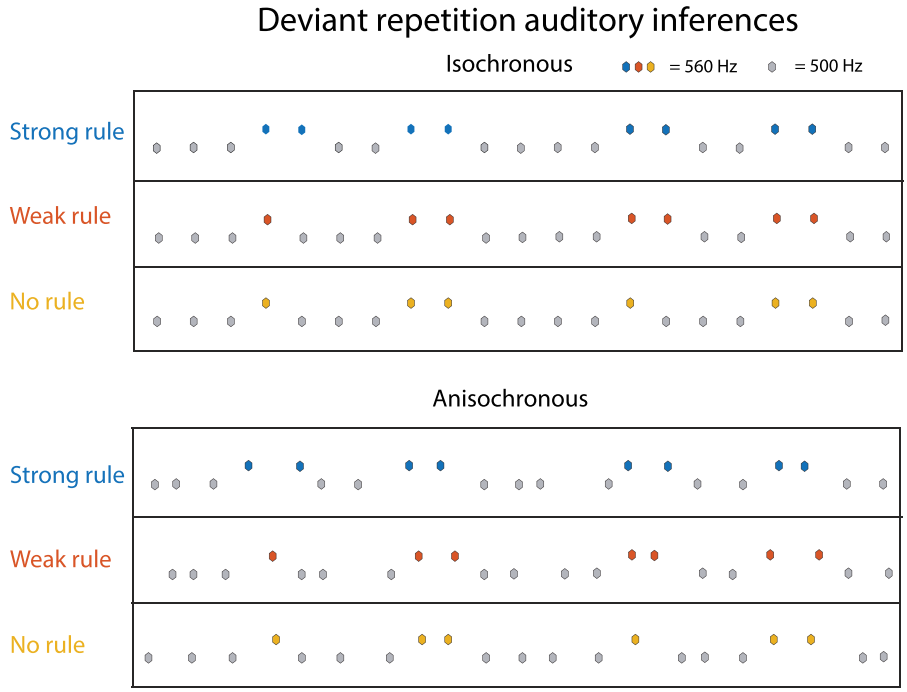

FIGURE 1 Sound sequences. Two 50-ms pure tones, a 500-Hz standard and a 560-Hz deviant, were arranged in isochronous or anisochronous sound sequences. Anisochrony was obtained using a ±20% jitter, uniformly distributed in steps of 5 ms for theta rhythms (6.67 Hz, corresponding to 150-ms stimulus onset asynchrony [SOA]), and 30 ms for delta rhythms (1.67 Hz, corresponding to 600-ms SOA). Blue colour codes for strong inference conditions (100% deviant repetition), red colour for weak inference conditions (75% deviant repetition) and yellow colour for the absence of a perceptual inference on repeated deviant onset (50% deviant repetition)

evidence for a strong inferential rule. Our proposal also fits an alternative way of looking at entrainment as the 'process by which a system returns to synchrony after a perturbation' (Large & Jones, 1999, p. 127), under the assumption that predictability in deviant event onset suppresses the amount of sensory surprise generated by the event. To extend the spectrotemporal entrainment hypothesis beyond slow rhythms, we tested the effects of conditional inferencing and temporal regularity at both delta (1.67 Hz) and theta stimulation rates (6.67 Hz).

2 | MATERIALS AND METHOD

2.1 | Participants

Data were collected at the Wilhelm Wundt Institute for Psychology of Leipzig University (Germany). An ERP analysis of part of the dataset was previously published (Tavano et al., 2014). Out of 20 originally recorded participants, all of whom gave their written informed consent to the study, 15 were viable for the present analysis (8 F, age mean = 25, *SD* ±4 years, range 20–30 years). Data from four participants contained recording mishaps, and the results of a fifth participant were discarded during the analysis stage, as very few deviant epochs survived a delta threshold (see Section 2.3). Participants sat alone in an electrically shielded, sound-attenuated chamber. They were instructed to direct their attention to a silenced, subtitled movie on a computer screen (15 inches, distance = 120 cm), while sound trains were diotically presented to them (pre-attentive auditory stimulation).

2.2 | Stimuli and experimental design

Sounds were two 50-ms pure tones, a 500-Hz standard and a 560-Hz deviant (ramp 10 ms, Tukey window). Sound sequences were delivered at two stimulation rates: delta rhythms running at a mean frequency of 1.67 Hz (mean stimulus onset asynchrony [SOA] = 600 ms) and theta rhythms running at a mean frequency of 6.67 Hz (mean SOA = 150 ms). For each stimulation rate, two factors were independently manipulated: temporal regularity and conditional inference (see Figure 1a). The temporal regularity factor was organized into two levels: isochronous versus anisochronous delivery, with jitter values up to ±20% randomly selected from a pool of nine uniformly distributed steps increasing/decreasing by 30 ms for delta sequences and 5 ms for theta sequences. The conditional inference factor was organized by the probability of deviant tone repetition. While the onset of a single deviant or of the first deviant tone of a pair within sound sequences was always unpredictable, repeated deviant tone probability was blocked by condition: 100% probability (*strong inference* condition), 75% (*weak inference* condition) or 50% (*no inference* condition).

The number of deviant pairs ($N = 120$) was kept constant across conditions to control for neural refractoriness effects in deviant repetition. In the strong inference condition, there were 1200 standard tones (83.3%) and 240 deviant stimuli (16.7%, no single deviants). In the weak inference condition, there were 1550 standard tones (86.2%) and 280 deviant stimuli (15.3%, 40 single deviants). In the no inference condition, there were 2400 standard tones (87%) and 360 deviant stimuli (13%, 120 single deviants). There were at least two standards preceding each first or single deviant onset. Delta stimulus sequences were delivered in 20 blocks, 10 isochronous—three each for strong and weak inference conditions and four for the no inference condition—and 10 anisochronous. Theta stimulus sequences were delivered in six blocks, three isochronous—one for each inference condition—and three anisochronous. Condition order presentation was randomly interleaved within participants.

2.3 | EEG data collection and preprocessing

The electroencephalogram (EEG) was DC-recorded at 512-Hz sampling rate using an ActiveTwo amplifier system (BioSemi, Amsterdam, the Netherlands; biosemi.com), with a 32-electrode cap organized according to the 10/20 system, and additional electrodes over the left and right mastoids. Horizontal and vertical eye movements were monitored using electrodes placed below the outer canthi of both eyes and at the nasion. The reference was placed on the tip of the nose.

EEG data were processed using custom scripts in Matlab (mathworks.com). Channels with technical malfunction (range 1–4 in seven participants) were interpolated using a spherical spline. EEG signal was offline passband filtered .2–40 Hz (Kaiser window, Beta 5.6533, filter length = 4637 points). Electroocular artefacts were corrected via regression weights, individually calculated using standardized ocular movements recorded prior to the experiment (Schlögel et al., 2007). The plugins *sphspline0.2* and *firfilt1.6.2* for the Matlab toolbox EEGLAB (eeglab.org, Delorme & Makeig, 2004) were used for interpolation and filtering, respectively (version 14_1_2b).

The strength of entrainment was calculated as the EEG power at stimulation frequency, estimated by means

of the fast Fourier transform (FFT), using the Matlab function *fft.m*. As the length of the input vector determines FFT resolution, to grant a fair comparison, we trimmed the data to the same segment length in all conditions. For theta stimulus sequences, each block was cut to 110,000 points (about 214 s), the duration of the strong inference condition. For delta stimulus sequences, we selected the first three blocks of each condition, cut each of them to 110,000 points and then concatenated the three blocks. The pre-processing pipeline was common to both stimulation rate contexts: *Z* scoring (electrode wise), first-order detrending, Hann windowing, applying a complex modulus (Matlab function *abs.m*) to the analytic signal (FFT) to extract global power estimates for each electrode using the absolute value corrected for the Hann window loss of power (sqrt(1.5)) and normalizing by the length of the Hann window (calculated using the next power of two: 131,072 points, equal to the number of frequency points to calculate the discrete Fourier transform). For delta stimulus sequences, we used a moving FFT window, with half-window overlap.

Auditory deviance detection has been shown to rely on at least two neural generators, a frontal one and a supratemporal one (Deouell, 2007). Hence, we analysed entrainment at both frontal and posterolateral (mastoid) electrodes, which likely pick up activity from both generators, but with different sensitivity: The frontal one likely reflects top-down effects on deviancy, and the supratemporal one is more sensitive to purely sensory deviancy magnitude (Schröger, 1998). The tip-of-the-nose reference helps avoiding systematic differences between hemispheres. We selected frontal (F3, Fz and F4) and mastoid (M1 and M2) electrodes to separately analyse the responses of the two neural generators. Scalp topographies were obtained using the *topoplot* function from EEGLAB. Spectral profile plots were obtained by averaging power estimates separately across frontal and mastoid electrodes. Noise-referenced power estimates for delta stimulus sequences were calculated for each frequency point by subtracting the average of two samples to the left and two to the right of the stimulation frequency bin, beginning from the second sample to each side. For theta stimulus sequences, we used six points to each side, again beginning from the second sample to each side.

Single-trial phase and power analyses were run on deviant pair trials (delta trial epoch: 4 s, 2-s baseline; theta trial epoch: 2.5 s, 1-s baseline) in isochronous sequences surviving a delta threshold (120 μV) applied to the selected frontal and temporal electrodes and further equated in number across conditions (range *N* trials for delta stimulus sequences: 30–66; for theta stimulus sequences, 61–81). A Hilbert transform was applied to bandpassed filtered trials for Fz electrode only (delta sequences, 1.2–2.2 Hz; theta sequences, 6.2–7.2 Hz; Kaiser window, Beta 5.6533, filter length = 3093 points), obtaining individual phase time courses in radians (Matlab function *angle.m*) and corresponding power time courses in μV^2. Plotting was helped by the functions *ShadedErrorBar.m* (Campbell, 2021), *barwitherr.m* (Callaghan, 2021) and *distributionPlot.m* (Jonas, 2021).

2.4 | Statistical analyses

Noise-referenced global power values were subjected to a series of univariate repeated-measures analyses of variance (rmANOVAs) using the Matlab function *anovan.m* with factors temporal regularity (isochronous and anisochronous), conditional inference (strong, weak and no inference) and stimulation rate (delta, theta), separately for frontal and mastoid sites. Two-sided, one-sample *t* tests were used to verify the presence/absence of entrainment in each condition, while two-sided, two-sample *t* tests were used to compare individual conditions. False discovery rate (FDR) correction with *Q* value = .05 was applied in all cases of multiple comparisons: both the threshold *p* value and the uncorrected *p* value are reported (Genovese et al., 2002).

To evaluate whether conditional inferences affect entrainment from the start, we matched the mean number of stimuli between delta and theta sequences by proportionally reducing the duration of the FFT input vector for theta sequences to a quarter relative to delta sequences, given that theta sequences run four times as fast as delta sequences: delta, 110,000 points, corresponding to about 214 s; theta, 27,500 points, equalling about 53 s. Then, we further parametrized those durations by dividing them by half, and then by four, in the end obtaining three FFT input vector lengths for each stimulation rates: long, medium and short. The resulting power estimates were subjected to a rmANOVA with factors segment duration (three levels), conditional inference (three levels) and stimulation rate (delta and theta).

Phase precision was calculated across trials within a peristimulus interval (200 ms for delta sequences and 50 ms for theta sequences) as the inverse of circular variance (Berens, 2009). A cluster-based permutation test (Matlab function permutest.m, Gerber, 2021) was used to detect significant differences between phase time courses within each stimulation rate. The correlation coefficient (Matlab function *corrcoef.m*) was used to estimate the correlation between single-trial phase and single-trial power, as well as between single-trial phase and noise-referenced global power. The significance of correlation peaks was determined via resampling by randomizing phase distribution across trials for each time point,

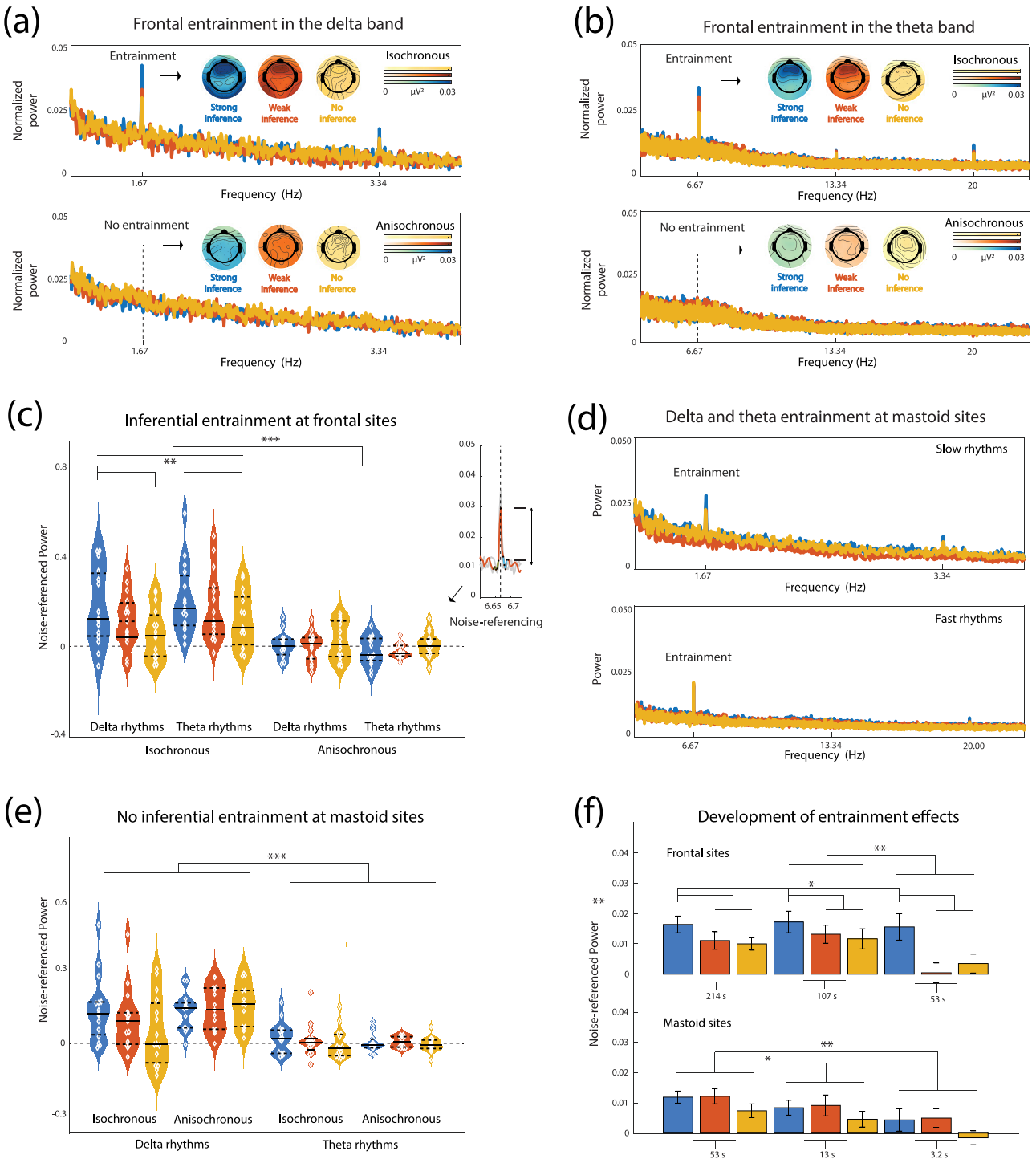

FIGURE 2 Legend on next page.

FIGURE 2 Temporal and spectral predictions co-modulate auditory entrainment. (a) Neural entrainment profiles were extracted from averaged frontal electrodes (F3, Fz and F4) activity. Scalp maps illustrate the spatial distribution of neural entrainment across conditions. Entrainment to isochronous stimulus sequences in the delta band was modulated by conditional inferencing, resulting in larger activity for the strong inference condition. The lower panel illustrates the absence of entrainment in all conditions when the stimulation is anisochronous. (b) Results in (a) were replicated for isochronous stimulus sequences in the theta band. (c) Noise-referenced power estimates highlighted a significant interaction between temporal regularity and conditional inference factors regardless of stimulation frequency. The inset to the right-hand side illustrates the noise-referencing procedure (cyan and green colours indicate the values to the right and left that are averaged and subtracted from the peak value). Violin plots highlight the shape of the distribution. Full horizontal bars indicate the median, while dashed ones indicate the interquartile range. (d) Entrainment resulting from averaged mastoid responses. (e) Entrainment at mastoid level shows only a main effect of temporal regularity. (f) The development of entrainment according to sensory evidence, matched by number of deviant pairs across stimulus frequencies. Frontal entrainment is independently modulated by sensory inference (strong inference) and the amount of sensory evidence for such inference. *p* value notation: **≤.01; *≤.05; ns = non-significant

creating 2000 random phase predictors per condition. Steiger's *Z* test was used to assess significant differences in correlation coefficient peaks. To determine the significance of peak latency differences, we used a jackknife resampling approach and entered the resulting estimates into a two-sided, two-sample *t* test.

3 | RESULTS

3.1 | Neural entrainment at frontal sites

The importance of regular onset time to generate entrainment was verified by a main effect of Temporal regularity: $F_{(1,28)} = 42.37$, $p \leq .0001$, $\eta p^2 = .60$. Anisochrony prevented entrainment in all conditions (FDR-corrected threshold $p = .0098$): All one-sample $ts_{(14)} \leq 2.09$, all p values $\geq .0544$ (see Figure 2a,b). Conversely, all isochronous conditions were well above noise level: all $ts_{(14)} \geq 4.26$, all $ps \leq .0007$, (see Figure 2c), except for the delta rhythm, no inference condition, which did not survive FDR correction ($t_{(14)} = 2.63$, $p = .0196$).

Confirming our hypothesis, a significant interaction emerged between the factors conditional inference and temporal regularity: $F_{(2,28)} = 11.55$, $p = .0002$, $\eta p^2 = .44$. Additionally, we found a main effect of conditional inference for isochronous sequences: $F_{(2,28)} = 7.29$, $p = .0028$, $\eta p^2 = .34$. The conditional inference factor was not significant for anisochronous sequences: $F_{(2,28)} = 2.29$, $p = .0925$. Isochronous tone sequences governed by a strong inference rule (mean noise-referenced power, SNR = .019, $SD = .010$) led to a significantly more robust neural entrainment than tone sequences with a no inference rule (mean = .011, $SD = .009$): $t_{(14)} = 3.59$, $p = .0028$. No other differences survived FDR correction (FDR-corrected threshold $p = .0167$): all $ts \leq 2.31$, all $ps \geq .0361$. Entrainment at delta and theta rhythms was similarly modulated by conditional perceptual inferences.

3.2 | Neural entrainment at mastoid sites

Regular onset time distribution was the sole determiner of entrainment at mastoid sites: $F_{(1,28)} = 26.68$, $p = .0001$, $\eta p^2 = .48$. Anisochrony prevented entrainment in all conditions (FDR-corrected threshold $p = .0098$): all one-sample $ts \leq 1.36$, all p values $\geq .1949$ (see Figure 3a,b). All isochronous theta conditions were above noise level: all $ts_{(14)} \geq 6.39$, all $ps \leq .0001$ (see Figure 3c). The significance of isochronous delta conditions did not survive FDR correction: all $ts_{(14)} \geq 2.41$, all $ps \leq .0300$. No other significant main effect or interaction was found: all $Fs \leq 1.88$, all $ps \geq .1708$.

3.3 | Development of entrainment effects

Does conditional inferencing determine the strength of entrainment early on, or do its effects develops gradually? At frontal sites, we found main effects of factors conditional inference: $F_{(2,28)} = 6.41$, $p = .0051$, $\eta p^2 = .33$, and segment duration: $F_{(2,28)} = 5.84$, $p = .0076$, $\eta p^2 = .29$. No significant effect of stimulation rate (delta vs. theta) was found: $F_{(2,28)} = .01$, $p = .9065$. This confirms that conditional inferences effects applied similarly within delta and theta bands. There were no significant interactions: all $Fs \leq 1.39$, all $ps \geq .1481$. Hence, we infer that conditional inferences contribute to the very generation of noise-normalized entrainment and are not applied at a later processing stage.

Regardless of segment duration, strong inference conditions (mean = .016, $SD = .011$) led to more robust entrainment than both weak (mean = .008, $SD = .0075$) and no inference (mean = .008, $SD = .009$) conditions: all $ts_{(14)} \geq 2.92$, all $ps \leq .0110$. No significant difference was found between weak and no inference conditions: $t_{(14)} = -.05$, $p = .9592$ (see Figure 2f, upper panel).

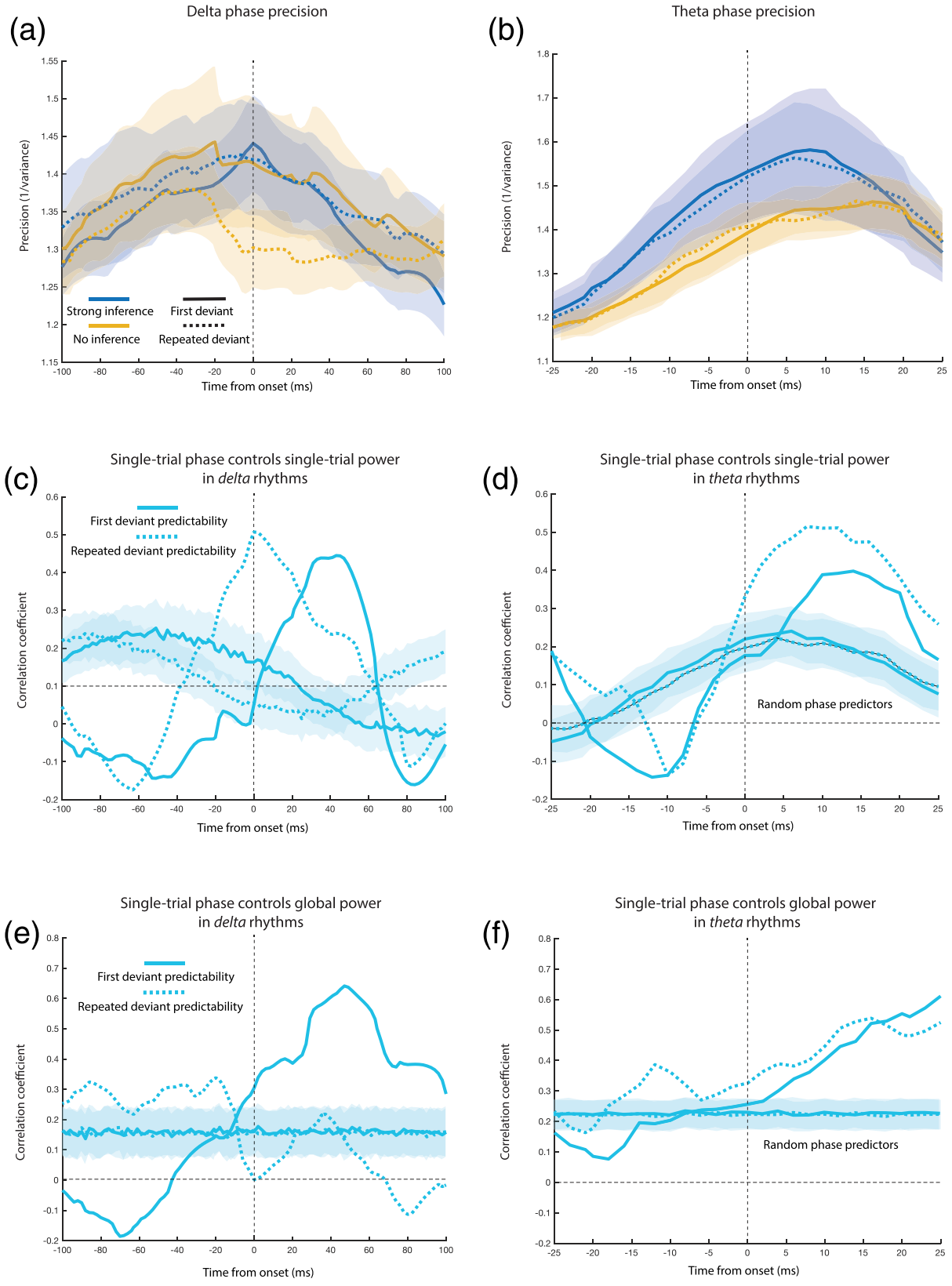

FIGURE 3 Legend on next page.

FIGURE 3 Phase precision governs entrained power. (a) Mean peristimulus phase precision across trials at 1.67 stimulation frequency. While the repeated deviant in the no inference condition displays diminished phase precision at onset, such difference did not reach statistical significance. Shaded areas indicate standard error of the mean boundaries. (b) Mean peristimulus phase precision across trials at 6.67 stimulus frequency. The apparent difference between strong and no inference conditions did not reach statistical significance. Shaded areas indicate standard error of the mean boundaries. (c) Predictability effects (strong minus no inference conditions) for both first and repeated deviants at 1.67 stimulus frequency. Phase precision differences positively correlate with single-trial power differences, suggesting that an increase in phase precision leads to an increase in power at stimulation frequency. Shaded areas reflect bootstrapping confidence intervals, obtained from random phase predictors. (d) Predictability effects (strong minus no inference conditions) for both first and repeated deviants at 6.67 stimulus frequency. Shaded areas reflect bootstrapping confidence intervals, obtained from random phase predictors. (e) Peristimulus phase concentration differences predict global power (full sequence) for both first and repeated deviants at 1.67 Hz. Notice that only repeated deviants are weakly but significantly correlated with global power in the prestimulus interval. (f) Peristimulus phase concentration differences predict global power (full sequence) for both first and repeated deviants at 6.67 Hz. Results are comparable with those at 1.67 Hz

Medium length segments (mean = .0140, SD = .0090) entrained more robustly than short segments (mean = .0065, SD = .0107): $t_{(14)} = 3.21$, $p = .0061$. Long segments (mean = .012, SD = .006) entrained more robustly than short segments, but the difference did not survive FDR correction: $t_{(14)} = 2.27$, $p = .0393$. Long and medium segments did not significantly differ: $t_{(14)} = -.78$, $p = .4477$.

At mastoid sites, entrainment was solely driven by the segment duration factor: $F_{(2,28)} = 21.62$, $p \leq .0001$, $\eta p^2 = .60$. Long segments (mean = .010, SD = .007) and medium segments (mean = .007, SD = .008) entrained more robustly than short segments (mean = .002, SD = .007): all $ts_{(14)} \geq 3.69$, all $ps \leq .0024$. In turn, long segments entrained more robustly than medium segments: $t_{(14)} = 2.72$, $p = .0163$. No other significant main effect of interaction was found: all Fs ≤ 3.20, all $ps \geq .0562$ (see Figure 2f, lower panel).

3.4 | Stimulus predictability controls the neural phase of entrained power

How do inferential rules modulate the contribution of deviant neural responses to auditory entrainment? We hypothesized that reducing the need to process the spectral features of predictable repeated deviants would significantly enhance the temporal precision in processing deviant onset across trials. To test this, we calculated peristimulus phase precision as the inverse of circular variance for both first and repeated deviant onset across individual trials, per participant and condition, separately for delta and theta sequences (see Figure 3a,b). A cluster-based permutation test highlighted no significant differences in either deviant position or conditional inference levels (strong vs. no inference conditions, the ones that showed significant differences), suggesting that phase precision was statistically similar across conditions: delta rhythms, all ts ≤ 20.38, all $ps \geq .1627$; theta rhythms, all ts ≤ 5.55, all $ps \geq .1336$.

However, while phase precision differences may not be statistically significant, they can still drive significant differences in their relationship with entrained power. To test this, we estimated the portion of the neural signal modulated by conditional inferencing by subtracting phase precision and mean single-trial power in the no inference conditions from phase precision and single-trial power in the strong inference conditions. If phase precision selectively governs entrained power under a strong inference rule, we should find positive correlations between phase precision and both single-trial power difference waves, reflecting local power control at deviant onset, as well as previously extracted FFT-based global power, reflecting the contribution of deviant tone phase to power estimates that also include standards. First, we correlated phase precision difference waves with mean single-trial power difference waves, for each participant and condition, at stimulation frequency. There resulted significant moderate positive correlations for both delta and theta rhythms, suggesting that an increase in phase precision also leads to more power at stimulation frequency: delta rhythms, first deviant peak correlation, rho = .42 (peak time = 47 ms), repeated deviant peak correlation, rho = .47 (right at deviant onset: peak time = 0); theta rhythms, first deviant peak correlation, rho = .39 (peak time = 14 ms), repeated deviant peak correlation, rho = .51 (peak time = 8 ms). The significance of correlation peaks was assessed by bootstrapping (2000 times) the calculation of the correlation coefficients: single-trial phase time courses were randomized before calculating phase precision differences across trials, separately for delta and theta rhythms as well as for first and repeated deviant (see Figure 3c,d). There was no significant effect of repetition in absolute correlation peak value for delta rhythmic sequences (Steiger's Z test Z-observed score = .156, $p = .8768$). However, for theta

rhythmic sequences, the effect was significant (*Z*-observed score = 3.29, $p = .0002$), suggesting that repetition increases the correlation of phase precision and mean single-trial power for predictable deviants. Furthermore, using a jackknife approach, we found that in delta rhythmic sequences, the repeated deviant correlation peaked significantly earlier than the first deviant correlation: $t(14) = 3.67$, $p = .0025$. This was also the case for theta rhythms, notwithstanding the small absolute difference in ms: $t(14) = 5.22$, $p = .0001$.

Finally, we calculated the correlation between phase precision difference waves and FFT-based global power differences between strong and no inference conditions. The significance of correlation peaks was again evaluated using a bootstrapping approach (2000 times), condition wise. There emerged significant post-onset, moderate-to-strong positive correlations for first deviants for delta sequences, rho = .64, $t = 4$ ms, and theta sequences, rho = .56, $t = 21$ ms, as well as for repeated deviants in theta sequences, rho = .54, $t = 15$ ms, and weak to moderate but significant pre-stimulus correlations for repeated deviants: delta, rho = .33, $t = -21$ ms; theta, rho = .39, $t = -12$ ms. We conclude that predictability, estimated as the difference between strong and no inference conditions, modulates deviant pair neural phases to positively contribute to global power at both delta and theta stimulation frequencies.

4 | DISCUSSION

Phase alignment of neural oscillations to rhythmical stimulus structures is a cornerstone paradigm of cognitive neuroscience (Lakatos et al., 2019), particularly in understanding how the brain encodes auditory sources (Zoefel & VanRullen, 2017). So far, phase alignment has been attributed to phase reset due to temporal regularity in event onset, although recent work suggests that, physiologically, neural entrainment as a mechanism for attention selection operates on both the temporal and spectral axes (Lakatos et al., 2013): attention suppresses the response of frequencies outside the stimulation frequency, while amplifying entrainment at stimulation frequency. It was however unclear if spectral predictability is computed even when attention is directed away from the stimulation frequency, and if so, by which mechanism it contributes to neural entrainment. In classic identical sound repetition paradigms, spectral predictability effects cannot be readily appreciated, because their import is not separable from that of temporal regularity. To overcome this limitation, we parameterized the effects of second-order predictability based on deviant event repetition probability and assumed that strongly inferring the spectral content of the next deviant event would perturb the temporal encoding of sound onset to a lesser degree, overall increasing sensory entrainment. Indeed, when sound structures favoured a strong perceptual inference—predicting with 100% probability that each deviant will be repeated—there resulted a significantly larger entrainment as compared with sound structures licencing no inference as to the spectral identity of the next sound (50% probability of deviant repetition). Hence, even within a pre-attentive setting, auditory entrainment at both slow and fast isochronous stimulation rates (1.67 and 6.67 Hz) appears to be concurrently modulated by temporal regularity and the statistical learning of spectral relationships between successive events. Such effect is visible at frontocentral sites, but not at mastoid sites, suggesting that supratemporal generators are mainly involved in tracking temporal regularity, while frontal ones are sensitive to both spectral and temporal predictability cues (Deouell, 2007; Schröger, 1998). Indeed, no matter the amount of evidence provided (short vs. long input segments), mastoid electrodes were sensitive only to the temporal component of entrainment, while frontal electrodes picked up spectral predictability cues right from the start, or with relatively limited evidence, along with temporal regularity.

Spectral predictability appears to operate through a phase control mechanism similar to the well-established drive of temporal regularity on concatenated phase resets governing phase alignment (Lakatos et al., 2019). Although peristimulus phase precision differences between conditions did not reach statistical significance per se, they modulated entrained power depending on both predictability and repetition. Repetition enhanced and/or anticipated the control exerted by cross-trial phase consistency on power at stimulation frequency. Interestingly, predictability affected the control on power exerted by phase precision not only for repeated deviants but also for first deviants. In previous work (Tavano et al., 2014), we showed that predictability affected only repeated deviants, using the deviant N1 neural event-related response (peaking between 100 and 150 ms post onset) as a dependent measure. Here, by investigating the peristimulus effects of phase concentration, we find that predictability changes the phase reset properties of both deviant events, even if the onset of first deviant events is, by distributional rule, unpredictable. This observation affords the novel perspective that spectral rules, once learned, become a property of, and affect the processing of, the entire stimulation sequence, rather than only of individual events within a sequence.

We found similar effects of predictability and repetition across delta (1.67 Hz) and theta (6.67 Hz) stimulation rates. This suggests the absence of a functional

division between delta and theta bands as far as the temporal and spectral dimensions of predictability in audition are concerned. The preference for the theta band that is well documented for neural speech encoding (Giraud & Poeppel, 2012; Teng & Poeppel, 2020) might stem from a processing bias that is stimulus specific (e.g., mean syllabic duration), rather than specific to the auditory modality. However, as a focus of analysis on the relationships between peristimulus phase and power is still relatively uncommon, future research is needed to tease apart the neural components of auditory predictability from those which specifically characterize important human stimuli, such as speech and music.

ACKNOWLEDGEMENTS

The authors wish to thank Nadin Greinert for collecting the original data, Dr. Lauren Fink of the Max Planck Institute for Empirical Aesthetics, Frankfurt am Main, for valuable feedback, and two anonymous reviewers. This project was made possible by a 'Reinhart Koselleck' grant of the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation), awarded to Erich Schröger.

CONFLICT OF INTEREST

The authors declare no competing financial interests.

AUTHOR CONTRIBUTIONS

AT devised research question. AT and BM analysed data. AT, BM and ES reviewed experimental results. AT drafted manuscript. AT, BM, DP and ES revised and approved manuscript.

PEER REVIEW

The peer review history for this article is available at https://publons.com/publon/10.1111/ejn.15630.

DATA AVAILABILITY STATEMENT

The pre-processed data that support the findings of this study are openly available in the Open Science Framework (OSF) repository, at https://osf.io/wf6dj/.

ORCID

Alessandro Tavano https://orcid.org/0000-0002-2096-5542
Burkhard Maess https://orcid.org/0000-0002-7857-291X
David Poeppel https://orcid.org/0000-0003-0184-163X
Erich Schröger https://orcid.org/0000-0002-8321-6629

REFERENCES

Ahmed, O. J., & Cash, S. S. (2013). Finding synchrony in the desynchronized EEG: The history and interpretation of gamma rhythms. *Frontiers in Integrative Neuroscience*, *7*, 58.

Barczak, A., O'Connell, M. N., McGinnis, T., Ross, D., Mowery, T., Falchier, A., & Lakatos, P. (2018). Top-down, contextual entrainment of neuronal oscillations in the auditory thalamocortical circuit. *Proceedings of the National Academy of Sciences (PNAS)*, *115*, E7605–E7614.

Berens, P. (2009). Circstat: A matlab toolbox for circular statistics. *Journal of Statistical Software*, *31*, 1–21.

Callaghan, M. Barwitherr. (https://www.mathworks.com/matlabcentral/fileexchange/30639-barwitherr-errors-varargin). MATLAB Central File Exchange (Retrieved November 15), 2021.

Campbell, R. (2021). raacampbell/shadedErrorBar (https://github.com/raacampbell/shadedErrorBar). GitHub (Retrieved November 15, 2021).

Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics. *Journal of Neuroscience Methods*, *134*, 9–21. https://doi.org/10.1016/j.jneumeth.2003.10.009

Deouell, L. (2007). The frontal generator of the mismatch negativity revisited. *Journal of Psychophysiology*, *21*, 188–203. https://doi.org/10.1027/0269-8803.21.34.188

Genovese, C. R., Lazar, N. A., & Nichols, T. (2002). Thresholding of statistical maps in functional neuroimaging using the false discovery rate. *NeuroImage*, *15*, 870–878. https://doi.org/10.1006/nimg.2001.1037

Gerber, E. M. Permutation test for dependent or independent measures of 1-D or 2-D data ([https://www.mathworks.com/matlabcentral/fileexchange/71737-permutest], matlab central file exchange.). MATLAB Central File Exchange (Retrieved November 15), 2021.

Giraud, A.-L., & Poeppel, D. (2012). Cortical oscillations and speech processing: Emerging computational principles and operations. *Nature Neuroscience*, *15*, 511–517. https://doi.org/10.1038/nn.3063

Jonas. Violin plots for plotting multiple distributions https://www.mathworks.com/matlabcentral/fileexchange/23661-violin-plots-for-plotting-multiple-distributions-distributionplot-m. MATLAB Central File Exchange (Retrieved November 15), 2021.

Keitel, A., & Gross, J. (2016). Individual human brain areas can be identified from their characteristic spectral activation fingerprints. *PLoS Biology*, *14*, e1002498. https://doi.org/10.1371/journal.pbio.1002498

Keitel, C., Quigley, C., & Philipp, R. (2014). Stimulus-driven brain oscillations in the alpha range: Entrainment of intrinsic rhythms or frequency-following response? *The Journal of Neuroscience*, *34*, 10137–10140. https://doi.org/10.1523/JNEUROSCI.1904-14.2014

Lakatos, P., Gross, J., & Thut, G. (2019). A new unifying account of the roles of neuronal entrainment. *Current Biology*, *29*, R890–R905. https://doi.org/10.1016/j.cub.2019.07.075

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*, 110–113. https://doi.org/10.1126/science.1154735

Lakatos, P., Musacchia, G., O'Connel, M. N., Falchier, A. Y., Javitt, D. C., & Schroeder, C. E. (2013). The spectrotemporal filter mechanism of auditory selective attention. *Neuron*, *77*, 750–761. https://doi.org/10.1016/j.neuron.2012.11.034

Lakatos, P., Shah, A. S., Knuth, K. H., Ulbert, I., Karmos, G., & Schroeder, C. E. (2005). An oscillatory hierarchy controlling

neuronal excitability and stimulus processing in the auditory cortex. *Journal of Neurophysiology*, *94*, 1904–1911. https://doi.org/10.1152/jn.00263.2005

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

Obleser, J., & Kayser, C. (2019). Neural entrainment and attentional selection in the listening brain. *Trends in Cognitive Sciences*, *23*, 913–926. https://doi.org/10.1016/j.tics.2019.08.004

O'Connell, M. N., Barczak, A., Schroeder, C. E., & Lakatos, P. (2014). Layer specific sharpening of frequency tuning by selective attention in primary auditory cortex. *Journal of Neuroscience*, *34*, 16496–16508. https://doi.org/10.1523/JNEUROSCI.2055-14.2014

Rossion, B. (2014). Understanding face perception by means of human electrophysiology. *Trends in Cognitive Sciences*, *18*, 310–318. https://doi.org/10.1016/j.tics.2014.02.013

Schlögel, A., Keinrath, C., Zimmermann, D., Scherer, R., Leeb, R., & Pfurtscheller, G. (2007). A fully automated correction method of EOG artifacts in EEG recordings. *Clinical Neurophysiology*, *118*, 98–104. https://doi.org/10.1016/j.clinph.2006.09.003

Schröger, E. (1998). Measurement and interpretation of the mismatch negativity. *Behavior Research Methods, Instruments, & Computers*, *30*, 131–145. https://doi.org/10.3758/BF03209423

Tavano, A., Widmann, A., Bendixen, A., Trujillo-Barreto, N., & Schröger, E. (2014). Temporal regularity facilitates higher-order sensory predictions in fast auditory sequences. *European Journal of Neuroscience*, *39*, 308–318. https://doi.org/10.1111/ejn.12404

Teng, X., & Poeppel, D. (2020). Theta and gamma bands encode acoustic dynamics over wide-ranging timescales. *Cerebral Cortex*, *30*, 2600–2614. https://doi.org/10.1093/cercor/bhz263

Zoefel, B., & VanRullen, R. (2017). Oscillatory mechanisms of stimulus processing and selection in the visual and auditory systems: State-of-the-art, speculations and suggestions. *Frontiers in Neuroscience*, *11*, 296.

How to cite this article: Tavano, A., Maess, B., Poeppel, D., & Schröger, E. (2022). Neural entrainment via perceptual inferences. *European Journal of Neuroscience*, 1–11. https://doi.org/10.1111/ejn.15630