

Neural Microstates Govern Perception of Auditory Input without Rhythmic Structure

Molly J. Henry,¹ Björn Herrmann,¹ and  Jonas Obleser^{1,2}

¹Max Planck “Auditory Cognition” Research Group, Max Planck Institute for Human Cognitive and Brain Sciences, 04103 Leipzig, Germany and

²Department of Psychology, University of Lübeck, 23562 Lübeck, Germany

Human perception fluctuates with the phase of neural oscillations in the presence of environmental rhythmic structure by which neural oscillations become entrained. However, in the absence of predictability afforded by rhythmic structure, we hypothesize that the neural dynamical states associated with optimal psychophysical performance are more complex than what has been described previously for rhythmic stimuli. The current electroencephalography study characterized the brain dynamics associated with optimal detection of gaps embedded in narrow-band acoustic noise stimuli lacking low-frequency rhythmic structure. Optimal gap detection was associated with three spectrotemporally distinct delta-governed neural microstates. Individual microstates were characterized by unique instantaneous combinations of neural phase in the delta, theta, and alpha frequency bands. Critically, gap detection was not predictable from local fluctuations in stimulus acoustics. The current results suggest that, in the absence of rhythmic structure to entrain neural oscillations, good performance hinges on complex neural states that vary from moment to moment.

Key words: auditory perception; entrainment; neural microstates; neural oscillations; psychophysics

Significance Statement

Our ability to hear faint sounds fluctuates together with slow brain activity that synchronizes with environmental rhythms. However, it is so far not known how brain activity at different time scales might interact to influence perception when there is no rhythm with which brain activity can synchronize. Here, we used electroencephalography to measure brain activity while participants listened for short silences that interrupted ongoing noise. We examined brain activity in three different frequency bands: delta, theta, and alpha. Participants' ability to detect gaps depended on different numbers of frequency bands—sometimes one, sometimes two, and sometimes three—at different times. Changes in the number of frequency bands that predict perception are a hallmark of a complex neural system.

Introduction

The neural system's capability to rapidly adapt to changing contexts is underpinned by its capacity to pliantly form and dissolve neural cell assemblies (Breakspear et al., 2004). The necessity for flexible neural (re)organization is likely especially valuable when the sensory context is unpredictable, as when stimuli lack rhythmic structure. We hypothesize that the dynamics of a neural sys-

tem operating in the absence of an organizing stimulus rhythm will be characterized by more complex contingencies between the phase of low-frequency neural activity and psychophysical performance than described previously for scenarios where entraining rhythms increased predictability and simplified neural dynamics.

Rhythmic structure is ubiquitous in behaviorally relevant environmental stimuli (Patel, 2008) and confers advantages such as enhanced detection or discrimination of temporally predictable relative to unpredictable target stimuli (Jones et al., 2002; McAuley and Jones, 2003; Rohenkohl et al., 2012; Henry and Herrmann, 2014; Lawrance et al., 2014.) The hypothesized mechanism underlying this behavioral advantage is entrainment of neural oscillations by rhythmic stimulus structure (Bishop, 1933; Lindsley, 1961; Buzsáki and Draguhn, 2004; Lakatos et al., 2005), which brings the excitable phase of the oscillation into line with temporally predictable stimulus events. Critically, in the context of complex rhythmic structure comprising two entraining frequencies, coupling between performance and neural phase is spe-

Received June 8, 2015; revised Nov. 2, 2015; accepted Dec. 2, 2015.

Author contributions: M.J.H., B.H., and J.O. designed research; M.J.H. performed research; M.J.H. and B.H. contributed unpublished reagents/analytic tools; M.J.H., B.H., and J.O. analyzed data; M.J.H., B.H., and J.O. wrote the paper.

This work was supported by a Max Planck Research Group grant from the Max Planck Society (J.O.). The authors declare no competing financial interests. The authors are indebted to Dunja Kunke, Christoph Daube, Steven Kalinke, and Elizabeth Mueche for help with data collection.

The authors declare no competing financial interests.

Correspondence should be addressed to Molly J. Henry, The Brain and Mind Institute, University of Western Ontario, London, ON, Canada. E-mail: mhenry55@uwo.ca.

DOI:10.1523/JNEUROSCI.2191-15.2016

Copyright © 2016 the authors 0270-6474/16/360860-12\$15.00/0

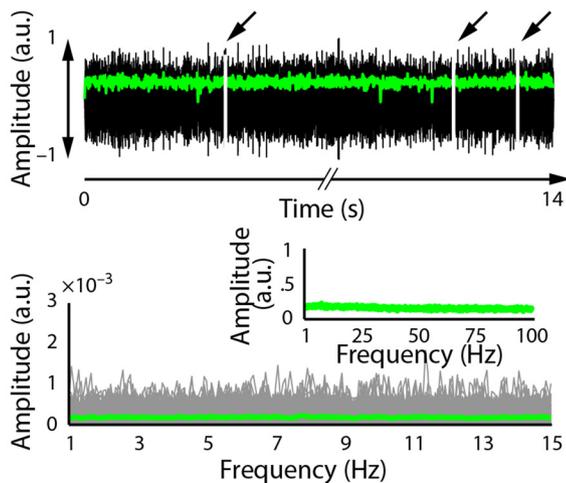
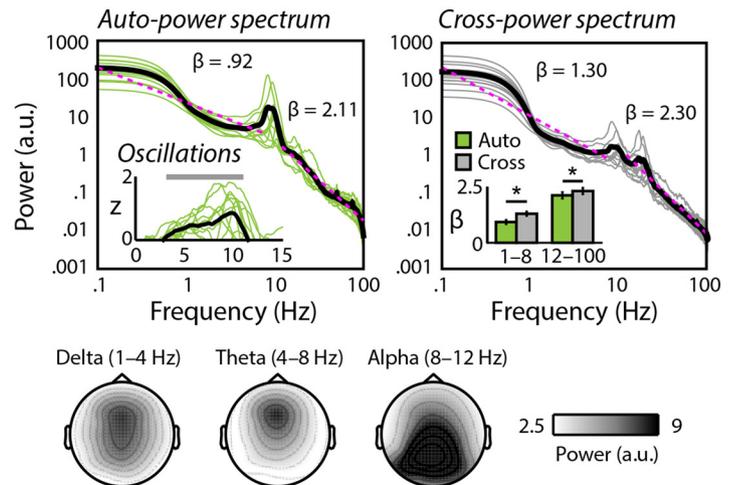
a Acoustic stimulus structure**b Frequency-domain neural responses**

Figure 1. *a*, Acoustic stimulus structure. Top, Stimuli were narrow-band noises without low-frequency rhythmic structure. Hilbert envelope is overlaid in green. Two, three, or four gaps (marked by arrows) were embedded in each 14 s stimulus. Bottom, Frequency-domain representation of acoustic stimuli for a single representative participant; amplitude spectra for individual stimuli are shown in gray, and the average over stimuli is overlaid in green. Inset, Average amplitude spectrum up to 100 Hz. *b*, Frequency-domain neural responses. Top, Autopower (left, both oscillatory and $1/f^\beta$ contributions) and cross-power (right, $1/f^\beta$ contributions only) spectral densities calculated for coarse-graining spectral analysis, plotted in log–log coordinates for single participants (green, gray) with the mean overlaid (black, thick line). Fits from which power-law exponents, β , were estimated for low (0.1–8 Hz) and high (12–100 Hz) frequencies are overlaid in pink. Subtracting cross-power spectral density from autopower spectral density yields an estimate of the oscillatory contributions to power (inset, left). Oscillatory power was significantly nonzero at all frequency bins between 3.4 and 10.75 Hz (marked by gray bar). Power-law exponents, β , differed significantly between auto- and cross-power spectral densities for both low ($p < 0.001$) and high ($p = 0.003$) frequencies (inset, right). Topographies are shown below for the classically defined delta (1–4 Hz), theta (4–8 Hz), and alpha (8–12 Hz) frequency bands in the autopower spectral density function.

cific to the entrained frequency bands and is not observed in frequency bands unrelated to the stimulus rhythm (Henry et al., 2014). Hence, neural entrainment by rhythmic structure would seem to reduce the dimensionality of complex neural dynamics (Freeman et al., 2001; Singer, 2013).

However, not all behaviorally relevant sensory input is well structured in time. In the absence of rhythmic structure, perception is hypothesized to proceed in a so-called “continuous,” or vigilance, mode (Schroeder and Lakatos, 2009; Schroeder et al., 2010; Henry and Herrmann, 2012), during which low-frequency neural activity should have a reduced impact on performance because of an overall suppression of power (Schroeder and Lakatos, 2009). Several studies have observed modulation of psychophysical performance by neural phase in the absence of rhythmic structure (Busch et al., 2009; Busch and VanRullen, 2010; Ng et al., 2012). Nonetheless, it remains an open question whether and how neural phase interacts across frequency bands to affect perception when auditory stimuli are not rhythmically structured and thus unpredictable.

We hypothesized that without rhythmic structure to increase stimulus predictability and potentially reduce the dimensionality of the neural state space via entrainment, neural phase effects would likely depend on higher-order interactions between phases in multiple frequency bands that change from moment to moment. The current human electroencephalography (EEG) experiment probed the complex neural states (here referred to as microstates) that lead to optimal performance in the absence of environmental rhythmic structure. We applied a novel analysis approach that considered instantaneous neural phase in multiple frequency bands and in different numbers of frequency bands at different times. We describe three neural microstates that yield optimal psychophysical performance and reflect a signature of a complex dynamic neural system.

Materials and Methods**Participants**

Eleven native German speakers (six female) with self-reported normal hearing took part in the study. All were right handed, and the mean age was 25.0 years (SD, 2.5 years). Data for one additional participant were discarded because of technical problems with the EEG recording. All participants gave written informed consent and received financial compensation. The procedure was approved of by the ethics committee of the medical faculty of the University of Leipzig and in accordance with the declaration of Helsinki.

Stimuli

Auditory stimuli were generated by MATLAB software (MathWorks) at a sampling rate of 44,100 Hz. Stimuli were 14 s narrow-band noises without low-frequency rhythmic structure (Fig. 1*a*). The center frequency of the complex carrier signals was randomized from trial to trial and took on values between 800 and 1200 Hz. All stimuli comprised 30 frequency components sampled from a uniform distribution with a 500 Hz range around the center frequency. The amplitude of each component decreased linearly with increasing distance to the center frequency. All stimuli were normalized with respect to peak amplitude and were presented at a comfortable level (~ 60 dB SPL).

Two, three, or four near-threshold gaps were inserted into each 14 s stimulus (gap onset and offset were gated with half-cosine ramps) without changing the duration of the stimulus. Gaps never occurred in the first or final 1.5 s of the 14 s stimulus and were constrained to fall no closer to each other than 1.5 s.

Procedure

Gap duration was first titrated for each individual listener such that detection performance was centered on 50%; the median individual threshold gap duration was 8.4 ms [semi-interquartile range (sIQR), ± 1.8 ms]. For the main experiment, the electroencephalogram was recorded while listeners detected gaps embedded in 14 s long narrow-band noise stimuli by pressing a button when they detected a gap. Listeners performed the task with eyes open while fixating a central cross. Overall,

each listener heard a total of 120 stimuli (three listeners initially heard 160, but the last 40 were discarded during subsequent processing); gaps were considered hits when the listener responded within 1.5 s of a gap presentation. Each listener was presented with a total of 480 gaps. The experiment lasted about 3 h, including preparation of the recording setup.

Data acquisition and analysis

The electroencephalogram was recorded from 26 Ag–AgCl scalp electrodes mounted on a custom-made cap (Electro-Cap International), according to the modified 10–20 system, and additionally from the left and right mastoids. Bipolar horizontal and vertical electrooculograms (EOGs) were also recorded. Signals were recorded continuously with a passband of DC to 135 Hz and digitized at a sampling rate of 500 Hz (TMS International). The on-line reference was the nose, and the ground electrode was placed at the sternum. Electrode resistance was kept under 5 k Ω . All EEG data were analyzed offline using custom MATLAB scripts and FieldTrip software (Oostenveld et al., 2011).

Raw data were epoched from 1.5 s before stimulus onset to 15.5 s following stimulus onset to capture the response to the full 14 s stimulus. Epoched data were then low-pass filtered at 100 Hz (zero-phase, sixth-order Butterworth) and rereferenced to linked mastoids. Blinks, muscle activity, electrical heart activity, and noisy electrodes were removed from the signal using independent component analysis (ICA) using the FieldTrip-implemented runica method (Makeig et al., 1996), which performs ICA decomposition using the logistic infomax algorithm (Bell and Sejnowski, 1995) with principle component dimension reduction. We demonstrated previously that the ICA routine implemented here does not bias estimates of neural phase (Henry et al., 2014). Individual trials were subsequently removed if the amplitude range exceeded 120 μ V; of the 120 presented trials, the median number of rejected trials was 10 (± 9.4 sIQR). Following artifact rejection, full stimulus epochs were analyzed in the frequency domain, and shorter epochs were defined that ranged between -2 s and $+2$ s with respect to each gap onset. Further analysis of pre-gap phase effects is described in Pre-gap phase effects, below.

Analysis of temporal structure in neural activity

Although the current stimuli did not contain low-frequency rhythmic structure (Fig. 1A) by which low-frequency neural activity could have been entrained, we nonetheless examined neural responses in the frequency domain to test whether there was low-frequency rhythmic (oscillatory) structure in the neural responses. For this purpose, we made use of coarse-graining spectral analysis (Yamamoto and Hughson, 1991; He et al., 2010). Neural responses corresponding to $+1$ to $+14$ s with respect to the auditory stimulation were considered so that onset- and offset-evoked responses would not be included in the spectral analysis.

Coarse-graining spectral analysis takes advantage of the self-similar nature of time series with $1/f^\beta$ structure (Mandelbrot and Van Ness, 1968) to separate $1/f^\beta$ (i.e., arrhythmic) and oscillatory contributions to total power. Two time series are constructed, one from the first half of the total time points, T , in the original time series $[x(t)]_{t=1, 2, 3, \dots, T/2}$, and a coarse-grained time series from every second time point in the original time series $[x'(t)]_{t=1, 3, 5, \dots, T}$. First, the total power of the neural signal (comprising both $1/f^\beta$ and oscillatory activity) is calculated as the autopower spectral density of $x(t)$. Then, the cross-power spectral density is calculated between $x(t)$ and $x'(t)$. The cross-power spectral density is assumed to reflect only power of $1/f^\beta$ noise. Subtracting the cross-power spectral density from the autopower spectral density yields the $1/f^\beta$ -free spectrum of $X(t)$, that is, the oscillatory contributions to the power spectrum (Yamamoto and Hughson, 1991).

We performed coarse-graining spectral analysis on single-trial time series data at each electrode; we report results for electrode Cz. Autopower and cross-power spectral densities were calculated using MATLAB's *cpsd* function, which returns the cross-power spectral density estimate of two time series using Welch's modified periodogram method. The power spectral densities were averaged over eight overlapping (50% overlap) Hamming-windowed sections of the 13 s neural response.

Since it was not clear the extent to which subtracting the cross-power

from the autopower spectral density might leave residual, chance-level power not attributable to neural oscillations, we tested significance by comparing against simulated data. For each participant, we simulated 1000 time-domain signals with $1/f^\beta$ structure [$\beta = 2$, as in the study by He et al. (2010)]. For each signal, we performed coarse graining. To make the power of the simulated signal directly comparable to the power of the neural signal, both differences (cross-power minus autopower) were normalized by their sum (cross-power plus autopower). Then, for each frequency bin between 0 and 15 Hz, we calculated a z score for the normalized difference from the neural signal with respect to the distribution of simulated signals. We then tested the resulting z scores against zero across participants using pointwise single-sample t tests. The p values were adjusted using a cluster-based correction for multiple comparisons (Maris and Oostenveld, 2007). The t statistics were summed within each "cluster" of consecutive frequency points. Then, on each of 1000 iterations, the labels identifying trials as having z scores or zero values were shuffled, and the t test and summing of values within consecutive clusters was repeated. A permutation distribution was formed from the maximum summed t statistic on each iteration. Summed t values from the original data set were considered significant when they were more extreme than the bottom 95% of the permutation values in the positive direction (corresponds to a one-tailed $\alpha = 0.05$).

We also tested for differences in estimates of the power-law exponent, β , between the autopower and cross-power spectral density functions. Power law exponents, β , were derived from linear fits to the auto- and cross-power spectral density functions plotted in log–log coordinates, separately for low frequencies (0.1 to 8 Hz) and high frequencies (12 to 100 Hz); we chose to discontinue the power-law fits in the alpha band because of pronounced peaks in this range that deviated from the roughly straight-line function of the power spectra plotted in log–log coordinates (Fig. 1b).

Pre-gap phase effects

To test for modulation of target (gap) detection performance due to prestimulus neural phase, we defined 4 s target epochs centered on the onset of each gap. We chose to perform pre-gap phase analyses only for electrode Cz for two reasons: First, the Cz electrode is located at the intersection of the topographies of spectral power for the delta, theta, and alpha frequency bands. Thus, the signal from this electrode reflects the combined contributions of delta, theta, and alpha generators. Second, and more practically, the computational demands of the present analysis prevented us from being able to perform the analysis for each sensor. However, we report topographies of the degree of optimal pre-gap phase clustering across participants (i.e., Rayleigh z values), which we used to define the microstates discussed here.

First, the single-trial time-domain signal was detrended (using linear regression), then gap-evoked responses were removed from the post-gap onset time window by multiplication with half of a Hann window that ranged between 0 and 50 ms after gap onset (and was zero thereafter). This was done to eliminate the possibility that "smearing" of the evoked response (or response differences between hit and miss trials) back into the prestimulus period by the wavelet convolution could produce spurious prestimulus phase effects (Lakatos et al., 2013; Henry et al., 2014). We conducted a simulation to confirm that this method is robust against phase distortions and found that application of the Hann window led to a very small phase negative phase shift (-0.22 rad on average) that was consistent across frequencies (between 1 and 15 Hz). Critically, the magnitude of the phase shift was identical for detected and undetected gaps, regardless of differences in postgap evoked responses. Thus, the method does not introduce a systematic phase difference between detected and undetected gaps that could have influenced the current results.

Next, the time-domain data were submitted to a wavelet convolution (wavelet width, 3 cycles) that yielded complex output in frequency bins ranging between 1 and 15 Hz (in 0.5 Hz steps) and with 2 ms temporal resolution (i.e., 500 Hz sampling frequency). Complex output was then converted to phase-angle time series. For each time–frequency bin in the range 1–15 Hz and for each time point in the window ranging from -0.5 to 0 s before gap, trials were sorted into 18 overlapping bins ranging

between $-\pi$ and $+\pi$ (bin width, 0.6π) based on single-trial phase values, and the hit rate was subsequently calculated for each bin. Then, the degree of behavioral modulation by the prestimulus phase was estimated in each time–frequency bin as the circular–linear correlation between the 18 phase angles and hit rates.

Circular–linear correlations in each time–frequency bin were converted to z scores based on a permutation procedure, in which single-trial behavioral responses were shuffled so that the relation to single-trial neural phase values was randomized. Trials were rebinned, and a circular–linear correlation was calculated; correlation values from 1000 iterations of this procedure made up a permutation distribution, from which a z score for the original circular–linear correlation was calculated. A time–frequency matrix (1 to 15 Hz, -0.5 to 0 s) of z scored circular–linear correlation coefficients was calculated for each participant. Then, a single-sample t test comparing the mean z score against 0 was calculated for each time–frequency bin using the FieldTrip-implemented cluster-based multiple-comparisons correction with a weighted cluster mean (wcm) criterion (wcm, 1.5) that takes into account both cluster size and the magnitude of test statistics within the cluster. In each resulting significant cluster, the consistency of the phase effects across participants was examined by testing the distribution of optimal neural phases (the phase where the hit rate was highest, estimated from cosine fits to individual participant data) against a uniform von Mises distribution using separate Rayleigh tests for each frequency band; p values were adjusted using a false-discovery-rate (FDR) correction for multiple comparisons (Benjamini and Hochberg, 1995; Benjamini and Yekutieli, 2001).

Pre-gap phase effects in delta-governed subsets of trials. To explore the possibility that pre-gap phase effects in the theta and alpha frequency bands might have depended on instantaneous delta states, we split the data into four overlapping bins (bin width, π) based on the delta phase. Separately for each delta bin, single trials in the theta, low-alpha, and high-alpha clusters were sorted into 18 overlapping phase bins as described above (bin width, 0.6π). Within each cluster, the consistency of the phase effects across participants was examined by testing the distribution of optimal neural phases (estimated from cosine fits to individual participant data) against a uniform von Mises distribution using separate Rayleigh tests for each delta bin; p values were adjusted using a false-discovery-rate correction for multiple comparisons (Benjamini and Hochberg, 1995; Benjamini and Yekutieli, 2001).

In the case that we observed a significant Rayleigh test in only one delta bin (R^*), we tested the resultant vector length for that bin against the mean resultant vector length averaged across the remaining delta bins (R^{ns}). Since the resultant vector reflects an across-participant measure, we opted for a permutation strategy to test statistical significance between resultant vectors. On each of 1000 iterations, for each participant and each of the four delta bins, binned hit rates were randomly rotated with respect to their corresponding phase values (across the 18 bins), keeping intact their relative ordering and nonindependence due to the use of overlapping bins. The optimal neural phase was estimated from a cosine fit to the permuted data. Then, the resultant vector length was calculated across participants for each delta bin. The mean resultant vector length across the R^{ns} bins was subtracted from the resultant vector length from the R^* bin. The result was a distribution of 1000 $R^* - R^{ns}$ differences against which we compared the actual $R^* - R^{ns}$ difference calculated from the real data. The z score and corresponding p value obtained by comparing the real value against the permutation distribution were taken as test statistics.

Higher-order interactive pre-gap phase effects in delta-governed subsets of trials. We also examined whether gap detection depended on specific phase–phase combinations in different frequency bands. Separately for each of four overlapping bins based on delta phase (bin width, π), we divided trials into a 50×50 grid of overlapping bins (bin width, 0.6π) for (1) theta and low-alpha phase, (2) theta and high-alpha phase, or (3) low and high-alpha phase. The hit rate grid was then smoothed with a two-dimensional filter (MATLAB's *imfilter*, modified to respect circular axes). The motivation for using a 50×50 grid in combination with smoothing for this analysis was because a cosine fit could not be used to estimate optimal phase for a two-dimensional data set. This is in contrast to the unidimensional phase analysis conducted separately for each fre-

quency band and described in the previous section. Instead, we used fine bin spacing in combination with two-dimensional smoothing to facilitate peak finding (i.e., locating the optimal combination of neural phases that yielded peak hit rate).

The neural phases of the two relevant frequencies corresponding to the peak gap-detection hit rate were retained. To statistically test the consistency of cross-frequency phase effects across participants, we developed a “spherical” Rayleigh test based on the equations provided by Leong and Carlile (1998; Zar, 2014); p values were adjusted using a false-discovery-rate correction for multiple comparisons across delta bins (Benjamini and Hochberg, 1995; Benjamini and Yekutieli, 2001).

To test the magnitude of the spherical resultant vector for a single significant delta bin (R^*) against the magnitudes of the vectors for the other bins (R^{ns}), we used a permutation strategy similar to that described above. Separately for each participant and delta bin, on each of 1000 iterations, we rotated the 50×50 grid of (unfiltered) hit rates with respect to their corresponding phase values in both frequency bands, keeping intact their relative ordering and nonindependence due to the use of overlapping bins. The resulting grid was then smoothed with a two-dimensional filter, the neural phases corresponding to peak hit rate were retained, and a spherical resultant vector length was calculated. The mean resultant vector length across the R^{ns} bins was then subtracted from the value from the R^* bin. The result was a distribution comprised of 1000 $R^* - R^{ns}$ differences, from which we determined a z score and p value for the actual $R^* - R^{ns}$ difference.

Time course analysis of neural microstates. To schematize the microstates we observed based on the above-described analyses, we simulated a neural oscillation starting from the time–frequency centroid of each significant cluster (delta, 1.75 Hz, -0.25 s; theta, 5.75 Hz, -0.22 s; low alpha, 10 Hz, -0.186 s; high alpha, 12 Hz, -0.016 s) based on the identified optimal phase (across-participant average). Then, for each of three neural microstates (i.e., different across-frequency phase combinations that resulted in peak hit rates at different times), we estimated the phase of the complex neural oscillation that would have optimally coincided with gap onset under idealized conditions. To test the accuracy of our extrapolations, we compared the predicted with the observed phase just before gap onset (the absolute duration in which we tested pre-gap phase depended on frequency and corresponded to 10% of a cycle). After adjusting the observed phase values for the -0.22 radian shift introduced by multiplication with half a Hann window to remove the ERP (see above, Pre-gap phase effects), we quantified the degree of inaccuracy of our extrapolations.

As a confirmation of the relation between the three neural microstates and gap-detection performance, we sorted each participant's single-trial time-domain data into three categories based on delta phase (corresponding to the three microstates; we ignored trials for which delta phase did not align with one of the microstates). We then filtered the time-domain data into between one and three frequency bands corresponding to the relevant frequency bands for each microstate, which we then summed. We tested the filtered time-domain signal preceding detected gaps (hits) against that preceding undetected gaps (misses) using paired-sample t tests at each time point and a cluster-based correction for multiple comparisons.

We were also interested in characterizing the time courses of the prevalence of each microstate as well as the time course of slow behavioral fluctuations. For each participant, we identified each single trial as belonging to one of three categories corresponding to the delta-governed microstates (again ignoring trials not corresponding to any of the microstates described here). Trials were categorized based on the phase(s) within the time–frequency windows identified in the previous analyses; phase bins were constructed individually for each participant around their own optimal phase in each frequency band (bin width, 0.6π). Then, we estimated microstate prevalence over time using a sliding window of 300 s duration moving in 5 s steps. The same was done for binary accuracy values. Finally, the time courses were transformed to the frequency domain using a fast Fourier transform after application of a Hann window.

Control analyses

We conducted two types of control analyses to rule out the possibilities that any observed neural phase effects on behavior might rather have been attributable to (1) modulation of behavior by pre-gap stimulus acoustics or (2) fluctuations in perceptual sensitivity coupled to eye movements.

Stimulus acoustics. It was suggested previously that neural phase effects in the absence of rhythmic structure specifically in audition are subject to a “third-variable problem,” whereby both neural activity and behavior are coupled to fluctuations in stimulus acoustics (VanRullen et al., 2014). Thus, we conducted an additional behavioral experiment (no electroencephalogram was recorded) identical to that already described with the exception that all audio files were retained for detailed acoustic analyses. Ten individuals that were not involved in the EEG experiment participated (procedures for obtaining informed consent and financial compensation were identical to the EEG experiment).

We correlated cochlea-scaled entropy (Stilp and Kluender, 2010) before gap occurrence with target detection performance in a time window corresponding to that in which we analyzed our neural phase data (i.e., for 0.5 s before gap occurrence). Cochlea-scaled spectral entropy is a measure of the relative unpredictability of acoustic signals, operationalized as the extent to which successive “spectral slices” differ from each other. Cochlea-scaled entropy (in particular, the replacement or deletion of high-entropy portions of an acoustic signal) has been shown to predict speech intelligibility better than the replacement or deletion of vowel segments, consonant segments, or simply segments of varying absolute duration (Stilp and Kluender, 2010); that is, deletions of high-entropy portions of a speech signal are more noticeable than deletions of relatively low-entropy portions of the signal. Thus, we expected that pre-gap cochlea-scaled spectral entropy might be related to gap-detection performance.

Cochlea-scaled entropy was calculated similarly to the description in Stilp and Kluender (2010). Each 16 ms “slice” of the stimulus was filtered by a bank of 12 symmetrical rounded-exponential filters with center frequencies equally spaced between 300 and 1700 Hz in equivalent-rectangular-bandwidth (ERB) units; filter bandwidths were four times the center frequency in ERBs (Patterson et al., 1982; Moore and Glasberg, 1983; Moore et al., 1990). The result was a 12 bin frequency histogram where the height of each bar was the amplitude of the corresponding filter output. Pairwise Euclidean distances were then calculated for frequency histograms belonging to neighboring slices. Finally, Euclidean distance values were averaged in a sliding rectangular window comprising five slices, resulting in a time course of cochlea-scaled entropy values.

For each participant, at each time point, single trials were binned into 18 overlapping percentile bins with 30% width based on cochlea-scaled entropy values. Gap-detection hit rates were calculated per bin and then correlated across bins with cochlea-scaled entropy percentile. Correlations were z transformed using a permutation strategy; on each of 1000 permutations, single-trial cochlea-scaled entropy values were shuffled with respect to binary accuracy values. Data were rebinned and correlated. At each time point, a z score was calculated for the actual correlation coefficient with respect to the mean and standard deviation of the correlations comprising the permutation distribution. Finally, z scores were tested against zero across participants using pointwise single-sample t tests. The p values were adjusted using a cluster-based correction for multiple comparisons (Maris and Oostenveld, 2007). The t statistics were summed within each “cluster” of consecutive time points. Then, on each of 1000 iterations, the labels identifying trials as having z scores or sign-flipped z scores were shuffled, and the t test and summing of values within consecutive clusters was repeated. A permutation distribution was formed from the maximum summed t statistic on each iteration. Summed t values from the original data set were considered significant when they were more extreme than the middle 95% of the permutation values in either direction (corresponds to a two-tailed $\alpha = 0.05$).

Eye movements. In light of recent observations that rhythmic motor activity benefits auditory perception (Su and Pöppel, 2012; Manning and Schutz, 2013; Morillon et al., 2014), we entertained the possibility that fluctuations in either (1) delta-band activity or (2) gap-detection performance might be related to the magnitude of eye movements before gap

onset. For both analyses, we made use of the horizontal and vertical EOG channels, which we high-pass filtered (0.9 Hz, 1099 points, Kaiser window), low-pass filtered (40 Hz, 93 points, Blackman window; Bosman et al., 2009), rereferenced to linked mastoids, and segmented into trials ranging from -0.5 to 0 s relative to gap onset. To compare to delta activity, we made use of data from electrode Cz. We high-pass filtered (0.9 Hz, 1099 points, Kaiser window), low-pass filtered (4 Hz, 701 points, Blackman window), rereferenced, and segmented into single trials. We only considered trials that had not been rejected for the analysis of the neural data.

For the analysis of delta-band activity, we correlated delta phase and the absolute magnitude of eye movements on a time point-by-time point basis. At each time point, we binned single trials based on single-trial delta phase (18 bins, 0.6π width, identical to our analysis of neural data) and averaged eye-movement magnitude within each bin. Then, we calculated a circular-linear correlation between delta phase and eye-movement magnitude. For the analysis of gap-detection performance, we binned single trials based on single-trial absolute eye-movement magnitude (separately for horizontal and vertical EOGs, 18 percentile bins, 30% width) and calculated a hit rate for each bin. Then, we calculated a linear correlation between eye-movement magnitude and hit rate. For both analyses, statistical testing and multiple-comparisons correction was performed following the same procedure as described for the analysis of cochlea-scaled entropy.

Effect size

Throughout this manuscript, effect sizes are reported as $r_{\text{equivalent}}$ (notated r_e throughout), which is equivalent to a Pearson product-moment correlation for two continuous variables, and to a point-biserial correlation for one continuous and one dichotomous variable (Rosenthal and Rubin, 2003). The only exception is for circular Rayleigh tests, where we report resultant vector length as the corresponding effect size measure (notated r for circular and R for spherical tests).

Results

Listeners detected near-threshold gaps embedded in 14 s narrow-band noise stimuli that did not possess rhythmic structure (Fig. 1a), in particular in the low frequencies in which neural phase has previously been linked to performance (Henry and Obleser, 2012; Neuling et al., 2012; Ng et al., 2012; Henry et al., 2014). The spectral content of the stimulus presented on each trial was unique, and we confirmed that local spectral content did not predict gap-detection performance. We probed for pre-gap neural microstates (i.e., combinations of phases in different frequency bands) that were associated with good gap-detection performance. Critically, we were interested in testing the possibility that neural phase in different frequency bands (and in different numbers of frequency bands) at different times might interact to determine auditory perception.

Behavioral performance

An adaptive-tracking procedure before the EEG experiment ensured that individually titrated gap durations resulted in detection performance near 50% (mean hit rate, 0.53 ± 0.048 SEM). Thus, the number of detected (hits) and undetected gaps (misses) was approximately equal.

Low-frequency neural activity reflects a combination of oscillatory and $1/f^\beta$ activity

Subtracting cross-power spectral density from autopower spectral density yields an estimate of the degree to which oscillations contribute to total power (Fig. 1b; Yamamoto and Hughson, 1991). Oscillatory strength was highest in the alpha band (at individually specific alpha frequencies), but oscillatory power was greater than zero at all frequencies in the range between 3.40 and 10.75 Hz ($p_{\text{cluster}} = 0.001$; we tested between 0 and 15 Hz). Thus,

neural activity below 3.40 Hz likely has an arrhythmic ($1/f^\beta$) temporal structure.

We estimated power-law exponents, β , separately for the autower and cross-power spectral density functions by fitting a linear function to the data in log–log coordinates. We did this separately for low (0.1–8 Hz) and high (12–100 Hz) frequencies (He et al., 2010); we chose to discontinue the power-law fits in the alpha band because of pronounced peaks in this range that deviated from the roughly straight-line function of the power spectra plotted in log–log coordinates (Fig. 1*b*). Power-law exponents, β , differed significantly between autower and cross-power spectral density for both low ($t_{(10)} = -13.57$, $p < 0.001$) and high ($t_{(10)} = -3.84$, $p = 0.003$) frequencies. Steeper slopes correspond to a faster drop off of power as a function of frequency for the cross-power spectral density function, and thus signify the potential presence of oscillatory contributions to total power.

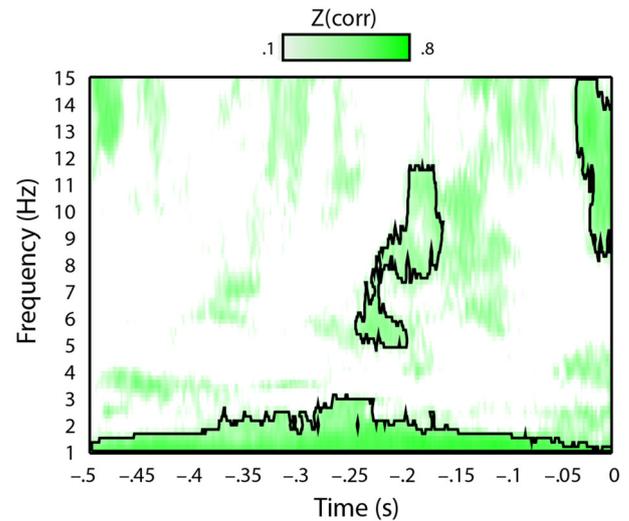
Neural phase effects on gap-detection performance

The relationship between neural phase and gap-detection performance was assessed by way of circular–linear correlations (Fig. 2*a*). Hit rates were calculated for each of 18 overlapping phase bins into which single trials were sorted, and correlations were then calculated between phase and hit rate. The statistical significance of z-scored circular–linear correlations was assessed for all frequencies ranging between 1 and 15 Hz and for time points preceding gaps by up to 500 ms at electrode Cz by means of nonparametric cluster statistics (Maris and Oostenveld, 2007). A significant relationship between hit rate and neural phase was observed for four time–frequency clusters: (1) delta band (1–2.5 Hz, -0.50 to 0 s, $r_e = 0.91$); (2) theta band (5–6.5 Hz, -0.24 to -0.20 s, $r_e = 0.74$); (3) low-alpha band (8–12 Hz, -0.23 to -0.16 s, $r_e = 0.69$); (4) high-alpha band (9–15 Hz, -0.03 to 0 s, $r_e = 0.88$); cluster p values < 0.05 (two-tailed). Figure 2*a* illustrates these four clusters.

To characterize in detail the nature of the effects in the significant time–frequency clusters, we estimated single-trial neural phases in time windows approximately centered in time within each cluster and chosen so that phase estimates were based on 10% of the cycle length in that frequency band (delta, 1–2.5 Hz, -0.28 to -0.22 s; theta, 5–6.5 Hz, -0.24 to -0.21 Hz; low alpha, 8–12 Hz, -0.20 to -0.18 s; high alpha, 9–15 Hz, -0.03 to -0.01 s). Phase values were averaged over frequency and time within a cluster, and single-trial accuracy values were again sorted into 18 overlapping bins based on phase. Gap-detection hit rates are plotted as a function of neural phase bin in Figure 2*b* (left). Notably, separate Rayleigh tests for each frequency band indicated that optimal phase (i.e., neural phase in which hit rate was highest) was consistent across participants only for the delta band ($z = 4.16$, $p_{\text{FDR}} = 0.048$, $r = 0.65$). This was in contrast to the theta ($z = 0.49$, $p_{\text{FDR}} = 0.94$, $r = 0.22$), low-alpha ($z = 0.06$, $p_{\text{FDR}} = 0.94$, $r = 0.08$), and high-alpha ($z = 1.89$, $p_{\text{FDR}} = 0.30$, $r = 0.44$) frequency bands, where nonsignificant phase concentrations indicated that optimal neural phase was inconsistent across participants (Fig. 2*b*, middle; optimal neural phases were estimated from cosine fits to behavioral data).

From a physiological standpoint, based on a model in which neural phase reflects local cortical excitability, consistency of optimal neural phase across participants would be expected. However, it has become a conventional, albeit less stringent, method of illustrating phase effects to perform a *post hoc* realignment of behavioral data so that peak hit rates for each individual participant all coincide with the same arbitrary neural phase (Busch et al., 2009; Neuling et al., 2012; Ng et al., 2012; Riecke et al., 2015).

a Correlation between neural phase and hit rate



b Single-participant data in significant clusters

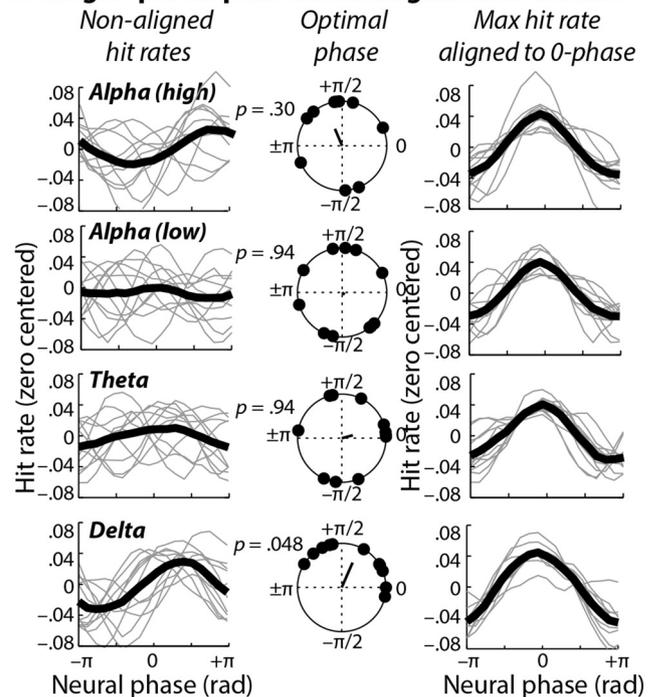


Figure 2. *a*, Correlation between neural phase and hit rate. Z-scored circular–linear correlations between binned gap-detection hit rates and neural phase for frequencies ranging between 1 and 15 Hz and for -0.5 to 0 s before gap, plotted for electrode Cz. Significant clusters (outlined in black) were observed in the delta (1–2.5 Hz, -0.50 to 0 s), theta (5–6.5 Hz, -0.24 to -0.20 s), low-alpha (8–12 Hz, -0.23 to -0.16 s), and high-alpha (9–15 Hz, -0.03 to 0 s) frequency bands. *b*, Single-participant data in significant clusters. Left, Nonaligned, zero-centered binned hit rates plotted as a function of neural phase in the significant delta (1–2.5 Hz), theta (5–6.5 Hz), low-alpha (8–12 Hz), and high-alpha (9–15 Hz) clusters; single-participant hit rates are shown in gray, and grand-average hit rates are overlaid in black. Middle, Analysis of optimal neural phase (i.e., the neural phase in which hit rate was highest) in each frequency band revealed that neural phase effects were only consistent across participants in the delta frequency band (delta, $p_{\text{FDR}} = 0.048$; all other $p_{\text{FDR}} \geq 0.30$). Right, Zero-centered binned hit rates plotted as a function of neural phase. Single-participant data were realigned so that maximum hit rates for each participant coincided (arbitrarily aligned with peak hit rate at phase = 0 rad); single-participant hit rates are shown in gray, and grand-average hit rates are overlaid in black.

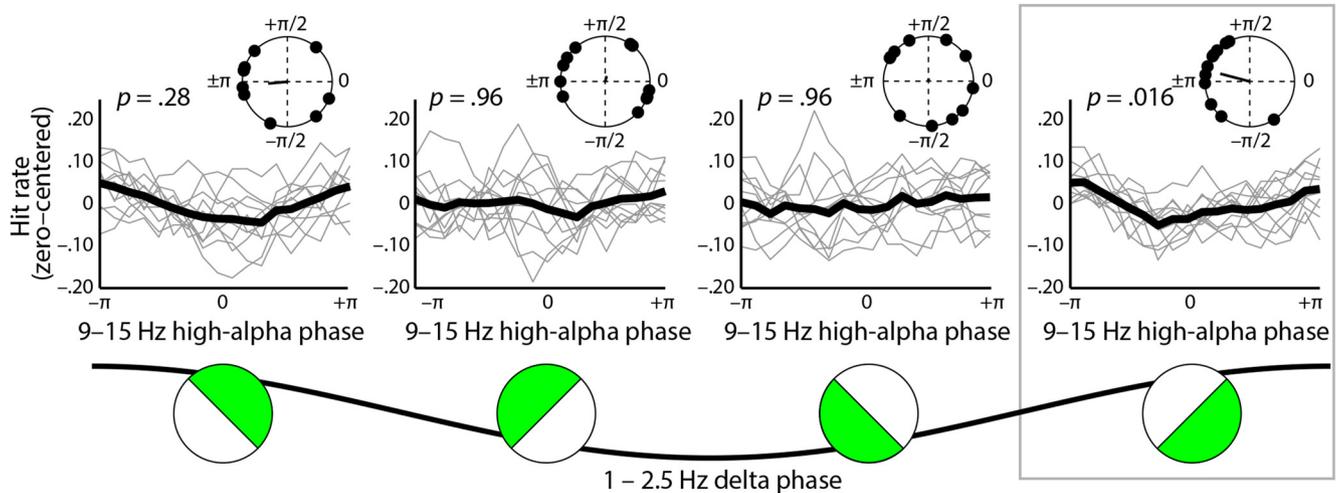


Figure 3. Delta phase mediated the influence of high-alpha phase on gap-detection hit rates. Hit rates (zero centered) are shown as a function of high-alpha (9–15 Hz) neural phase, separately for each of four delta phase bins (illustrated at bottom of figure). Single-participant data are shown in gray, and grand-average data are overlaid in black (thick lines). Circle plots show distributions of optimal neural phases estimated from cosine fits with resultant vectors. High-alpha (9–15 Hz) phase had a consistent influence on gap-detection hit rates only when delta (1–2.5 Hz) phase ranged between $\pi/4$ and $-3\pi/4$ (encompassing 0; far right).

Figure 2*b* (right) emulates this convention and plots gap-detection hit rates as a function of neural phase with single-participant data realigned so that peak hit rates coincide at phase = 0 rad. Plotted this way, relative phase effects can be discerned in all frequency bands under consideration here. However, we consider it important that only the delta-band phase effect is present without such single-participant data realignment and that best delta phase is thus consistent across participants also in absolute terms.

Delta phase determined higher-frequency phase effects

The previous analysis indicated that individual participants showed modulation of hit rates by neural theta and alpha phase. However, the specific neural phase that corresponded to best hit rate (optimal phase) was inconsistent across participants in both frequency bands. To further investigate this apparent inconsistency, we investigated the possibility that behavioral performance patterns were explainable by phase–phase interactions across frequency bands. In particular, we expected that an analysis based on delta phase would help clarify the theta- and alpha-dependent results.

First, we separated the single-trial data into four bins based on delta phase. Then, for each delta-phase bin, hit rates were calculated as a function of neural phase (based on sorted and binned single trials) independently for the theta, low-alpha, and high-alpha clusters (i.e., what we will refer to as first-order interactions). We estimated optimal neural phase per participant based on cosine fits to the hit rates, separately for theta, low-alpha, and high-alpha frequency bands (in each delta phase bin) and tested for consistency of optimal neural phase across participants using Rayleigh tests.

Interestingly, for those trials in which delta phase ranged between $\pi/4$ and $-3\pi/4$ (encompassing 0 radian), we observed a significant clustering of optimal high-alpha neural phases (yielding best hit rates) across participants ($z = 5.10$, $p_{\text{FDR}} = 0.016$, $r = 0.68$; Fig. 3). We did not observe such an emergence of a consistent phase effect in the high-alpha frequency band for any of the remaining three delta phase bins ($p_{\text{FDR}} \geq 0.28$). Moreover, a direct comparison revealed that high-alpha consistency (as indexed by resultant vector length) was indeed significantly

stronger for the delta bin ranging between $\pi/4$ and $-3\pi/4$ (encompassing 0 radian) compared to the average over the remaining delta bins ($z = 4.03$, $p < 0.001$). Finally, splitting trials based on delta phase did not improve the across-participant consistency of the influence of neural phase on hit rate in either the theta or low-alpha frequency bands ($p_{\text{FDR}} \geq 0.62$).

Of note, we confirmed that delta amplitude did not differ between delta-phase bins when we took into account the full 0.5 s before gap occurrence ($p = 0.93$), or when we looked specifically in the time windows of the significant clusters in the higher frequency bands (theta, $p = 0.90$; low alpha, $p = 0.94$; high alpha, $p = 0.83$).

Higher-order interactions between frequency bands determine performance

We next tested whether, for each of the four individual delta bins, gap-detection hit rates were predictable from a combination of phases in the theta, low-alpha, and high-alpha clusters (based on multidimensional binning of single trials; here referred to as higher-order interactions). Using a spherical Rayleigh test, we found that when delta phase ranged between $3\pi/4$ and $-\pi/4$ (encompassing π), hit rates were highest for a specific combination of theta and low-alpha phase that was marginally significantly clustered across participants ($z = 3.94$, $p_{\text{FDR}} = 0.06$, $r = 0.60$; Fig. 4). This delta phase bin was notably different from that for which we observed a consistent optimal high-alpha phase. For the remaining delta phase ranges, combinations of theta and low-alpha phase did not predict hit rates consistently across participants ($p_{\text{FDR}} \geq 0.28$). Moreover, the consistency of the theta–low-alpha interaction in the delta bin ranging between $3\pi/4$ and $-\pi/4$ (encompassing π) was significantly higher than the consistency of the theta–low-alpha phase effects averaged over the other delta bins ($z = 3.62$, $p < 0.001$). Finally, splitting the data based on delta phase did not reveal consistent combinatorial effects of either theta and high alpha or low alpha and high alpha ($p_{\text{FDR}} \geq 0.64$).

Gap-detection performance depends on delta-governed neural microstates

Figure 5*a* (insets) shows schematic characterizations of the different brain states leading to optimal gap-detection performance based on extrapolating the time course of prototype narrow-

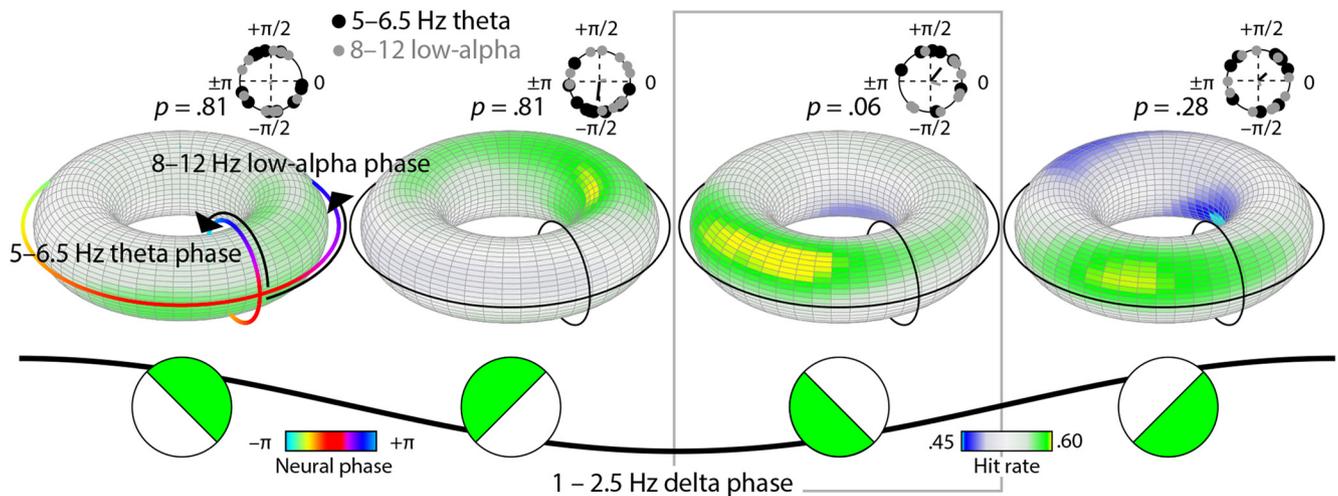


Figure 4. Delta phase mediated the joint influence of theta and low-alpha phase on gap-detection hit rates. Hit rates are plotted on tori as a function of 5–6.5 Hz theta phase (smaller, inside circle) and 8–12 Hz low-alpha phase (larger, outside circle), separately for each of four delta phase bins (illustrated at bottom of figure). Theta and low-alpha phase marginally significantly comodulated hit rates only for trials on which delta phase ranged between $3\pi/4$ and $-\pi/4$ (encompassing π), to an extent that was significantly greater than for all other delta bins. Circle plots (top) are provided for visualization purposes and show the neural phase yielding the peak hit rate separately for the theta and alpha frequency bands. p values correspond to spherical Rayleigh tests for theta and alpha phases considered together (FDR corrected).

band neural oscillations from each observed phase effect up to gap onset. These schematized time courses are shown together with single-participant and mean time courses of data identified as belonging to each microstate and filtered as such. The schematic microstates and time-domain data correspond well. For each frequency band, extrapolation inaccuracy was small (delta, 0.04 ± 0.03 radian, 95% confidence interval; theta, 0.08 ± 0.68 rad; low alpha, -0.47 ± 1.01 rad; high alpha, 0.05 ± 0.02 rad). Notably, for each of these three schematic microstates, best gap-detection performance occurred when gaps coincided with the local near trough-to-rising phase of the complex neural signal. For each microstate, we compared time courses of data preceding detected (hits) versus undetected (misses) gaps and found that these differed significantly in time ranges that corresponded to our initial circular-linear correlation analysis reported in Figure 2 (Microstate A, $p_{\text{cluster}} < 0.001$; Microstate B, $p_{\text{cluster}} \leq 0.003$; Microstate C, $p_{\text{cluster}} \leq 0.02$). In particular, time courses were more negative before gaps that were subsequently detected than for gaps that were undetected. Topographies showing the strength of clustering of optimal neural phase (i.e., Rayleigh z values) revealed distinct spatial distributions of each microstate.

Time courses of microstate prevalence are shown in Figure 5b for three representative participants. Frequency-domain representations revealed slow fluctuations in the prevalence of individual microstates as well as gap-detection performance that had an average rate of 0.003 Hz (~ 6 min) across participants (Fig. 5b, bottom).

Control analyses

We analyzed the relationship between gap-detection hit rates and stimulus acoustics to rule out the possibility that the dependence of hit rates on distinct neural microstates was attributable to variations in stimulus acoustics that drove both neural phase and behavioral fluctuations (for a similar analysis, see Ng et al., 2012). We did not observe a significant correlation between cochlea-scaled entropy and gap-detection hit rates at any pre-gap time point ($p_{\text{cluster}} \geq 0.07$, two-tailed; Fig. 6a).

We also examined the relation between eye-movement magnitude before gap onset and (1) delta phase as well as (2) gap-

detection hit rates in light of recent behavioral work indicating that movement has a positive effect on auditory perception (Su and Pöppel, 2012; Manning and Schutz, 2013; Morillon et al., 2014). Horizontal EOG magnitude did not correlate significantly with either delta phase ($p_{\text{cluster}} \geq 0.18$, $r_e \leq 0.42$) or gap-detection performance ($p_{\text{cluster}} = 0.10$, $r_e \leq 0.50$). Vertical EOG magnitude was not correlated with delta phase ($p_{\text{cluster}} \geq 0.22$, $r_e \leq 0.38$). However, pre-gap vertical EOG magnitude was significantly negatively correlated with gap-detection hit rates in a very brief time window (-0.474 to -0.466 s) that did not overlap with any of the neural phase effects ($p_{\text{cluster}} = 0.004$, $r_e = 0.76$; Fig. 6b). In contrast to observations of an auditory perception benefit associated with movement, gaps were more likely to be missed when vertical EOG magnitude was large. This is potentially unsurprising as the benefit of movement is specific to rhythmic movements in time with rhythmic stimuli (for a similar result on ERPs, see Schmidt-Kassow et al., 2013; Fautrelle et al., 2015).

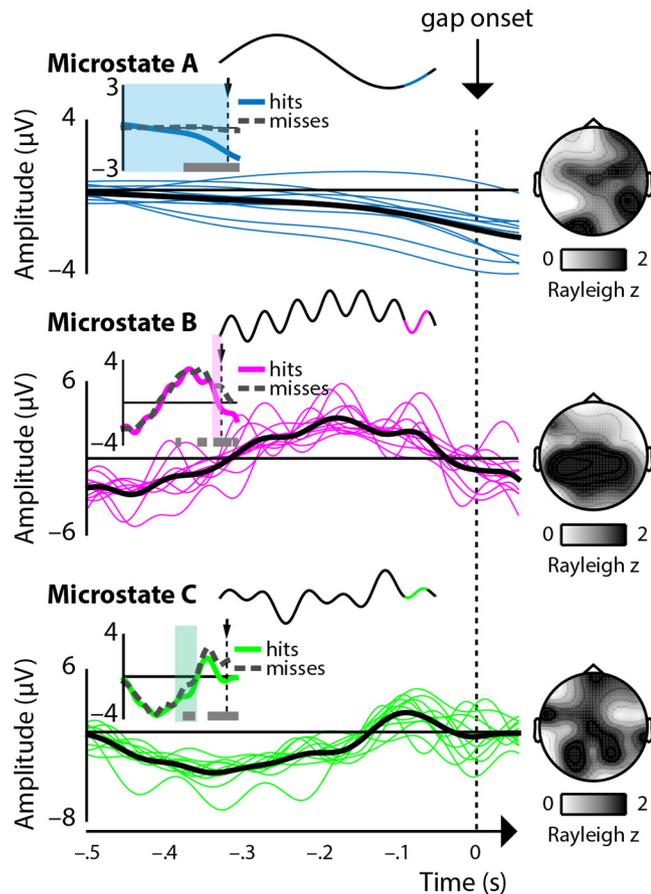
Discussion

Gap-detection hit rates peaked in three delta-governed microstates that were characterized by (1) delta phase alone, (2) a combination of delta and high-alpha phase, or (3) a combination of delta, theta, and low-alpha phase. The results demonstrate that in the absence of a regular stimulus rhythm to entrain neural activity, contingencies between brain states and behavior are characterized by spectrotemporal complexity.

Delta-governed neural microstates determined gap-detection performance

Gap detection depended on specific phase-phase combinations across frequency bands—sometimes one, sometimes two, and sometimes three frequency bands—that were relevant on different trials. First, gap detection depended on arrhythmic delta-band activity in the range of 1–2.5 Hz (a single-band effect). Second, neural oscillatory phase in a high-alpha (9–15 Hz) cluster modulated gap detection, but optimal high-alpha phase was only consistent across participants for a subset of data chosen based on delta phase (a first-order interaction). Finally, gap detection depended on a specific combination of oscillatory theta

a Distinct neural microstates



b Microstate time course

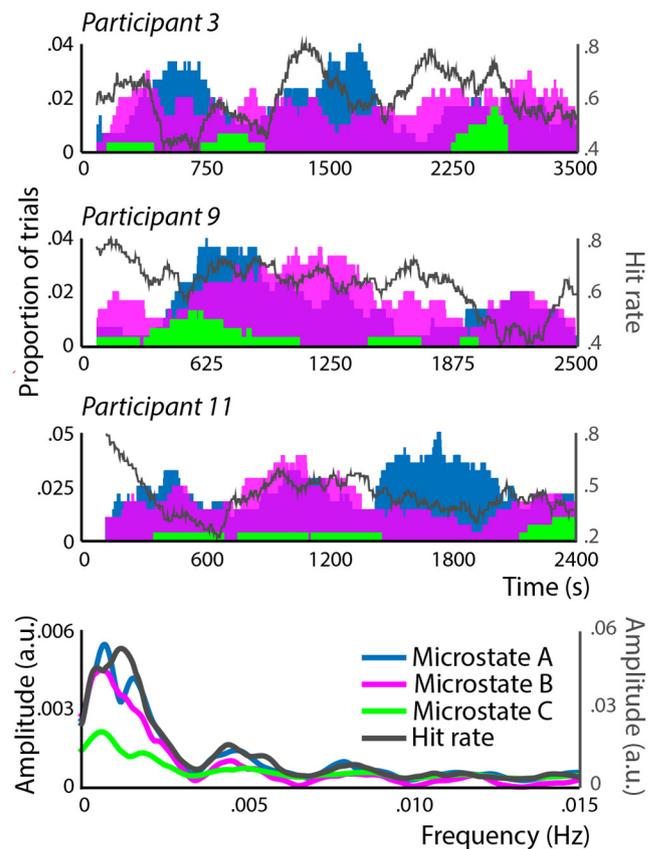


Figure 5. Gap-detection performance was determined by distinct delta-governed neural microstates. **a**, When hit rates were determined by delta phase alone (Microstate A), by a combination of delta and high-alpha phase (Microstate B), and by a combination of delta, theta, and low-alpha phase, gap-detection hit rates were relatively high when the local near trough-to-rising phase (highlighted in blue, pink, and green, respectively) of the complex neural activity coincided with gap onset (insets). Single-participant (thin, colored lines) and mean (thick, black lines) time course data corresponded well to the schematic microstates. Insets compare time courses preceding detected (hits, blue, pink, green) and undetected (misses, gray dashed lines) gaps. The time courses differed significantly (marked by gray bars below time courses) in time ranges corresponding to our analysis based on circular-linear correlations (Fig. 2, marked by colored shaded bars), and polarity was more negative preceding hits than misses. Topographies show the strength of consistency across participants in terms of optimal neural phase for gap-detection performance. **b**, Time courses of microstate prevalence (shown here for three representative participants) revealed slow fluctuations in the prevalence of individual microstates as well as gap-detection performance that had an average rate of 0.003 Hz across participants (bottom).

(5–6.5 Hz) and low-alpha (8–12 Hz) phase, but the optimal joint theta-low-alpha phase was only consistent across participants for a different subset of data again chosen based on delta phase (a higher-order interaction). The specific delta phase ranges and spatial topographies associated with peak performance were different for the three scenarios. Thus, auditory perception in the absence of rhythmic stimulus structure was determined by the presence of distinct delta-governed neural microstates characterized by combinations of arrhythmic ($1/f^\beta$) and oscillatory activity.

$1/f^\beta$ structure is ubiquitous in natural time series including brain activity, earth seismic activity, and financial markets (Kaiser and Ermentrout, 2010; He, 2014), although its functional neural significance is a matter of ongoing research. Neurally, $1/f^\beta$ activity signifies the presence of long-range temporal correlations, and as such reflects a “memory” for system dynamics (Linkenkaer-Hansen et al., 2001), which is disrupted in Alzheimer’s disease (Montez et al., 2009), schizophrenia (Nikulin et al., 2012), and major depressive disorder (Linkenkaer-Hansen et al., 2005). $1/f^\beta$ power has been demonstrated to correlate with population firing rate (Ray and Maunsell, 2011) and to show task

specificity (Palva et al., 2013). The current results indicate that $1/f^\beta$ activity interacts with neural oscillatory activity to play a functional role in auditory perception.

Modeling work suggests that the phase of neural activity at different time scales in the delta-theta-alpha frequency range may play somewhat different roles in encoding sound. Electrophysiological recordings from auditory cortex of anesthetized rats indicate that multiunit activity is modulated by phase-phase combinations in the 1–12 Hz range in the presence of auditory stimulation and in silence (Kayser et al., 2015). Modeling indicated that delta-band phase modulates background firing, whereas higher-frequency phase simultaneously modulates response gain. Thus, we hypothesize that our delta-governed neural microstates may reflect simultaneous modulation of background neuronal firing by delta phase and modulation of response gain at faster (theta and alpha) time scales.

The current approach can be contrasted with previous research in which lack of optimal neural phase consistency was remedied by realigning data so that individual participants’ behavioral peaks coincided at an arbitrary phase (Busch et al., 2009; Busch and VanRullen, 2010; Neuling et al., 2012; Ng et al., 2012;

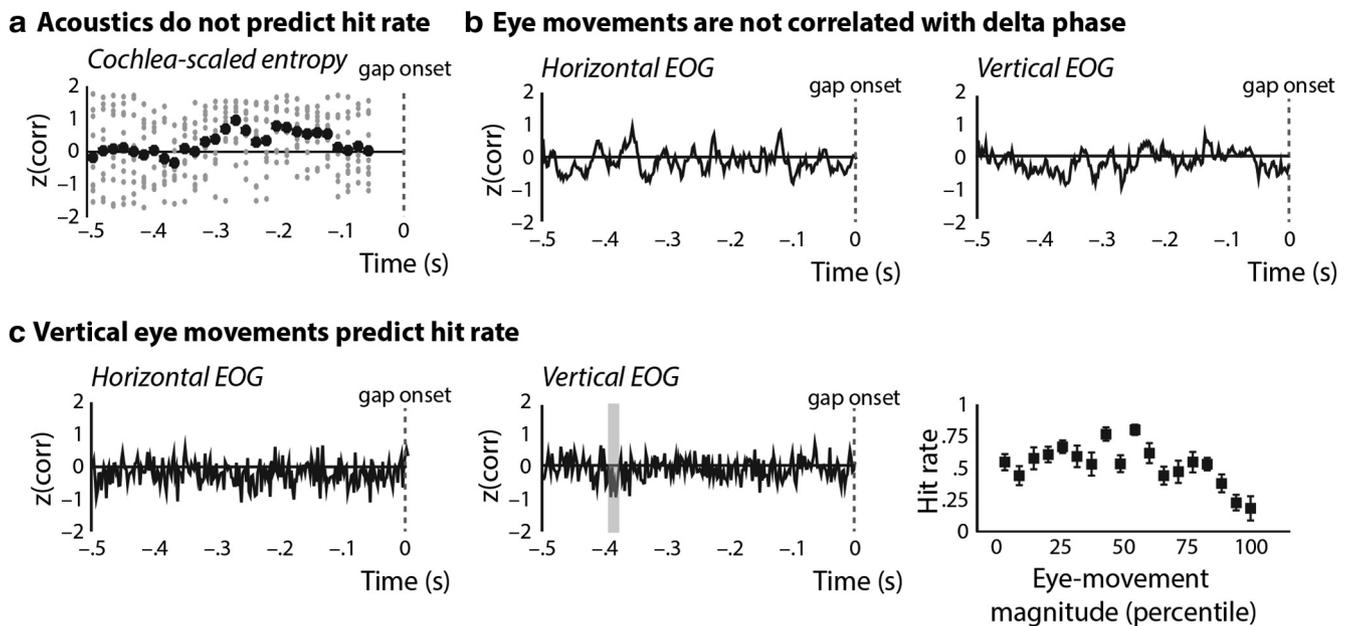


Figure 6. *a*, Cochlea-scaled entropy was not significantly correlated with gap-detection hit rates (y -axis shows z -transformed correlation coefficients as a function of pre-gap time); gray dots show single-participant data, and black dots show mean data. *b*, Neither horizontal (left) nor vertical (right) EOG activity was correlated with delta phase. *c*, Horizontal EOG activity (left) was not correlated with hit rate, but vertical EOG (middle) activity was negatively correlated with gap-detection hit rates in a brief time window that did not overlap with the neural phase effects (highlighted gray). The right panel shows hit rates as a function of vertical EOG eye-movement magnitude (percentile, mean \pm SEM).

Riecke et al., 2015; Fig. 2*b*). Notably, the majority of studies that have failed to observe across-participant consistency in optimal neural phase (Busch et al., 2009; Busch and VanRullen, 2010; Ng et al., 2012) also used rhythmically unpredictable stimuli. In contrast, in our own previous work using rhythmic stimuli, we observed significant across-participant consistency in the distribution of optimal neural phases in specific frequency bands corresponding to the stimulus rhythm (Henry and Obleser, 2012; Henry et al., 2014). We suggest that in the absence of rhythmic structure at a specific frequency, across-participant consistency in the distribution of optimal neural phases may not be expected in a single frequency band, but may only be apparent in subsets of trials corresponding to distinct neural states.

We did not observe a consistent relationship between gap-detection performance and pre-gap stimulus acoustics (cochlea-scaled entropy; see also Ng et al., 2012). More generally, converging evidence suggests that neural phase effects in auditory perception cannot result simply from acoustic confounds. First, we showed previously that neural phase is a better predictor of performance than stimulus phase (Henry and Obleser, 2012). Second, auditory target detection depends on the phase of neural activity entrained by transcranial direct current stimulation in the absence of an acoustic rhythm (Neuling et al., 2012). Finally, we consider it unlikely that the complex pattern of neural phase effects observed in the current study would be simply attributable to acoustic fluctuations. In sum, auditory perception in the absence of low-frequency rhythmic structure depends on delta-governed neural microstates that cannot be explained by acoustic features of the stimulation.

Microstates in neural time and space

It is notable that neural complexity and neural microstates are often defined in spatial terms (Lehmann et al., 1994; Kondakor et al., 1997; Müller et al., 2005; Britz et al., 2009; Sporns, 2014). Spatial complexity, considered crucial for cognitive function (Tononi et al., 1994), is defined in terms of a balance between

unique patterns of activity within individual brain regions and simultaneous communication between brain regions (Friston et al., 1995; Tononi et al., 1998; Sporns et al., 2000). Empirically, the degree of spatial network complexity measured before the occurrence of a near-threshold target predicts detection performance for tactile stimulation (Weisz et al., 2014). Similarly, changes in patterns of between-region connectivity predict the percept associated with an ambiguous visual stimulus (Hipp et al., 2011). Future work, integrating local dynamic spectrotemporal with global spatiotemporal considerations of neural complexity and neural microstates, will be necessary to mechanistically characterize the neural preconditions for human perception, cognition, and consciousness (Tononi and Edelman, 1998).

Neural microstates reveal complex neural dynamics in the absence of rhythmic structure

The perspective adopted here is rooted in an approach to neuroscience that sees the brain as a complex dynamic system. On this view, nonlinearities are expected to arise in spatial, spectral, and temporal aspects of neural processing because of the neurophysiological, pharmacological, and histological properties of the brain and neural tissue (Freeman et al., 2001; Breakspear et al., 2004). Here, spectrotemporal nonlinearities were revealed, whereby the relation between neural phase and behavior, and the number and identity of relevant frequency bands, changed over time. We suggest that the presence of distinct neural microstates associated with good gap-detection performance reflects flexible reorganization of brain dynamics over time (Freeman et al., 2001); that is, the neural states that were associated with good performance at different times were characterized by interactions between different numbers of frequency bands, suggesting moment-to-moment variations in the dimensionality of the neural states or shifts between stable attractors in a high-dimensional neural state space.

The facility to switch flexibly between neural state configurations, to uncouple and recouple functional neural assemblies, is

critical to ensure that the neural system is capable of rapid adaptation to changing contexts (Breakspear et al., 2004). This is likely especially true when sensory contexts are unpredictable, as when stimuli lack rhythmic structure. We hypothesize that the association of distinct microstates with good performance constitutes a hallmark of a neural system operating in the absence of an organizing stimulus like an auditory rhythm that increases predictability and reduces complexity of brain dynamics through entrainment. We consider it important for future research to characterize rhythm's role in shaping neural dynamics and the dimensionality thereof.

We propose that the special role of slow, arrhythmic delta activity in determining the microstates is related to the overall weakly stable nature of the neural system in the absence of rhythmic structure (Friston, 2000; Breakspear, 2004). Flexible state reorganizations would not be adaptive on a very fast time scale. Thus, we speculate that such state transitions could be reflected in slow delta-band neural activity, leading to the observation of what we have termed delta-governed microstates. Interestingly, this hypothesis and indeed the current data can be contrasted with the idea that in the absence of rhythmic structure to entrain low-frequency neural activity, the importance of the delta band should be minimized, mainly by suppression of delta-band activity altogether (Schroeder and Lakatos, 2009). Our data do not provide evidence for such a suppression and, together with evidence of the importance of delta-band activity in silence (Kayser et al., 2015), emphasize the sustained influence of delta phase on performance in so-called “continuous-mode processing.”

Conclusions

The present EEG study presents a novel analysis approach that considers higher-order contingencies between frequency bands to explain behavior. The results demonstrate that the phase of neural activity affects target-detection performance in the absence of low-frequency rhythmic structure (during continuous-mode processing). Gap detection was best, and optimal neural phase was consistent across participants, in three distinct, delta-governed neural microstates in which the number and identity of frequency band(s) relevant for predicting performance varied over time. The results suggest that in the absence of predictability afforded by rhythmic stimulus structure, optimal detection performance hinges on moment-to-moment variations neural state organization.

References

- Bell AJ, Sejnowski TJ (1995) An information maximization approach to blind separation and blind deconvolution. *Neural Comput* 7:1129–1159. [CrossRef Medline](#)
- Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc B Met* 57:289–300.
- Benjamini Y, Yekutieli D (2001) The control of the false discovery rate in multiple testing under dependency. *Ann Stat* 29:1165–1188. [CrossRef](#)
- Bishop GH (1933) Cyclic changes in the excitability of the optic pathway of the rabbit. *Am J Physiol* 103:213–224.
- Bosman CA, Womelsdorf T, Desimone R, Fries P (2009) A microsaccadic rhythm modulates gamma-band synchronization behavior. *J Neurosci* 29:9471–9480. [CrossRef Medline](#)
- Breakspear M (2004) “Dynamic” connectivity in neural systems. *Neuroinformatics* 2:205–226. [CrossRef Medline](#)
- Breakspear M, Williams LM, Stam CJ (2004) A novel method for the topographic analysis of neural activity reveals formation and dissolution of ‘dynamic cell assemblies.’ *J Comput Neurosci* 16:49–68. [Medline](#)
- Britz J, Landis T, Michel CM (2009) Right parietal brain activity precedes perceptual alternation of bistable stimuli. *Cereb Cortex* 19:55–65. [CrossRef Medline](#)
- Busch NA, VanRullen R (2010) Spontaneous EEG oscillations reveal periodic sampling of visual attention. *Proc Natl Acad Sci U S A* 109:16048–16053. [Medline](#)
- Busch NA, Dubois J, VanRullen R (2009) The phase of ongoing EEG oscillations predicts visual perception. *J Neurosci* 29:7869–7876. [CrossRef Medline](#)
- Buzsáki G, Draguhn A (2004) Neuronal oscillations in cortical networks. *Science* 25:1926–1929. [Medline](#)
- Fautrelle L, Mareschal D, French R, Addyman C, Thomas E (2015) Motor activity improves temporal expectancy. *PLoS One* 10:e0119187. [CrossRef Medline](#)
- Freeman WJ, Kozma R, Werbos PJ (2001) Biocomplexity: Adaptive behavior in complex stochastic systems. *Biosystems* 59:109–123. [CrossRef Medline](#)
- Friston KJ (2000) The labile brain ii: transients, complexity and selection. *Philos T R Soc Lond B Biol Sci* 355:237–252. [CrossRef](#)
- Friston KJ, Tononi G, Sporns O, Edelman GM (1995) Characterizing the complexity of neuronal interactions. *Hum Brain Mapp* 3:302–314. [CrossRef](#)
- He BJ (2014) Scale-free brain activity: past, present, and future. *Trends Cogn Sci* 18:480–487. [CrossRef Medline](#)
- He BJ, Zempel JM, Snyder AZ, Raichle ME (2010) The temporal structures and functional significance of scale-free brain activity. *Neuron* 66:353–369. [CrossRef Medline](#)
- Henry MJ, Herrmann B (2012) A precluding role of low-frequency oscillations for auditory perception in a continuous processing mode. *J Neurosci* 32:17525–17527. [CrossRef Medline](#)
- Henry MJ, Herrmann B (2014) Low-frequency neural oscillations support dynamic attending in temporal context. *Timing Time Percept* 2:62–86. [CrossRef](#)
- Henry MJ, Obleser J (2012) Frequency modulation entrains slow neural oscillations and optimizes human listening behavior. *Proc Natl Acad Sci U S A* 109:20095–20100. [CrossRef Medline](#)
- Henry MJ, Herrmann B, Obleser J (2014) Entrained neural oscillations in multiple frequency bands comodulate behavior. *Proc Natl Acad Sci U S A* 111:14935–14940. [CrossRef Medline](#)
- Hipp JF, Engel AK, Siegel M (2011) Oscillatory synchronization in large-scale cortical networks predicts perception. *Neuron* 69:387–396. [CrossRef Medline](#)
- Jones MR, Moynihan H, MacKenzie N, Puente J (2002) Temporal aspects of stimulus-driven attending in dynamic arrays. *Psychol Sci* 13:313–319. [CrossRef Medline](#)
- Kayser C, Ermentrout B (2010) Complex times for earthquakes, stocks, and the brain's activity. *Neuron* 13:329–331. [Medline](#)
- Kayser C, Wilson C, Safaai H, Sakata S, Panzeri S (2015) Rhythmic auditory cortex activity at multiple timescales shapes stimulus-response gain and background firing. *J Neurosci* 35:7750–7762. [CrossRef Medline](#)
- Kondakor I, Lehmann D, Michel CM, Brandeis D, Kochi K, Koenig T (1997) Prestimulus EEG microstates influence visual event-related potential microstates in field maps with 47 channels. *J Neural Transm* 104:161–173. [CrossRef Medline](#)
- Lakatos P, Shah AS, Knuth KH, Ulbert I, Karmos G, Schroeder CE (2005) An oscillatory hierarchy controlling neuronal excitability and stimulus processing in auditory cortex. *J Neurophysiol* 94:1904–1911. [CrossRef Medline](#)
- Lakatos P, Musacchia G, O'Connell MN, Falchier AY, Javitt DC, Schroeder CE (2013) The spectrotemporal filter mechanism of auditory selective attention. *Neuron* 77:750–761. [CrossRef Medline](#)
- Lawrance, ELA, Harper, NS, Cooke, JE and Schnupp, JWH (2014) Temporal predictability enhances auditory detection. *J Acoust Soc Am Express Letters* 135:EL357–EL363. [CrossRef](#)
- Lehmann D, Michel CM, Pal I, Pascual-Marqui RD (1994) Event-related potential maps depend on prestimulus brain electric microstate map. *Int J Neurosci* 74:239–248. [CrossRef Medline](#)
- Leong, P and Carlile, S (1998) Methods for spherical data analysis and visualization *J Neurosci Methods* 30:191–200.
- Lindsley, DB (1961) The reticular activating system and perceptual integration. (Sheer DE, ed). Austin, TX: Austin University of Texas.
- Linkenkaer-Hansen K, Nikouline VV, Palva JM, Ilmoniemi RJ (2001) Long-range temporal correlations and scaling behavior in human brain oscillations. *J Neurosci* 21:1370–1377. [Medline](#)
- Linkenkaer-Hansen K, Monto S, Ryttsälä H, Suominen K, Isometsä E, Käh-

- könen S (2005) Breakdown of long-range temporal correlations in theta oscillations in patients with major depressive disorder. *J Neurosci* 25:10131–10137. [CrossRef Medline](#)
- Makeig, S, Bell, AJ, Jung, TP and Sejnowski, TJ (1996) Independent component analysis of electroencephalographic data. (Touretzky D, Mozer M, Hasselmo M, eds). Cambridge, MA: MIT.
- Mandelbrot BB, Van Ness JW (1968) Fractional Brownian motions, fractional noises, and applications. *SIAM Rev* 10:422–437. [CrossRef](#)
- Manning F, Schutz M (2013) “Moving to the beat” improves timing perception. *Psychon Bull Rev* 20:1133–1139. [CrossRef Medline](#)
- Maris E, Oostenveld R (2007) Nonparametric statistical testing of EEG- and MEG-data. *J Neurosci Methods* 164:177–190. [CrossRef Medline](#)
- McAuley JD, Jones MR (2003) Modeling effects of rhythmic context on perceived duration: a comparison of interval and entrainment approaches to short-interval timing. *J Exp Psychol Hum Percept Perform* 29:1102–1125. [CrossRef Medline](#)
- Montez T, Poil SS, Jones BF, Manshanden I, Verbunt JP, van Dijk BW, Brusaard AB, van Ooyen A, Stam CJ, Scheltens P, Linkenkaer-Hansen K (2009) Altered temporal correlations in parietal alpha and prefrontal theta oscillations in early-stage alzheimer’s disease. *Proc Natl Acad Sci U S A* 106:1614–1619. [CrossRef Medline](#)
- Moore BC, Glasberg BR (1983) Suggested formulae for calculating auditory-filter bandwidths and excitation patterns. *J Acoust Soc Am* 74:750–753. [CrossRef Medline](#)
- Moore BC, Peters RW, Glasberg BR (1990) Auditory filter shapes at low center frequencies. *J Acoust Soc Am* 88:132–140. [CrossRef Medline](#)
- Morillon B, Schroeder CE, Wyart V (2014) Motor contributions to the temporal precision of auditory attention. *Nat Commun* 5:5255. [CrossRef Medline](#)
- Müller TJ, Koenig T, Wackermann J, Kalus P, Fallgatter A, Strik W, Lehmann D (2005) Subsecond changes of global brain state in illusory multistable motion perception. *J Neural Transm* 112:565–576. [CrossRef Medline](#)
- Neuling T, Rach S, Wagner S, Wolters CH, Herrmann CS (2012) Good vibrations: Oscillatory phase shapes perception. *Neuroimage* 63:771–778. [CrossRef Medline](#)
- Ng BS, Schroeder T, Kayser C (2012) A precluding but not ensuring role of entrained low-frequency oscillations for auditory perception. *J Neurosci* 32:12268–12276. [CrossRef Medline](#)
- Nikulin VV, Jönsson EG, Brismar T (2012) Attenuation of long-range temporal correlations in the amplitude dynamics of alpha and beta neuronal oscillations in patients with schizophrenia. *Neuroimage* 61:162–169. [CrossRef Medline](#)
- Oostenveld R, Fries P, Maris E, Schoffelen JM (2011) FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Lect Notes Comput Sci* 2011:1–9. [CrossRef](#)
- Palva JM, Zhigalov A, Hirvonen J, Korhonen O, Linkenkaer-Hansen K, Palva S (2013) Neuronal long-range temporal correlations and avalanche dynamics are correlated with behavioral scaling laws. *Proc Natl Acad Sci U S A* 110:3585–3590. [CrossRef Medline](#)
- Patel AD (2008) Music, language, and the brain. New York: Oxford UP.
- Patterson RD, Nimmo-Smith I, Weber DL, Milroy R (1982) The deterioration of hearing with age: Frequency selectivity, the critical ration, the audiogram, and speech threshold. *J Acoust Soc Am* 72:1788–1803. [CrossRef Medline](#)
- Ray S, Maunsell JH (2011) Different origins of gamma rhythm and high-gamma activity in macaque visual cortex. *PLoS Biol* 9:e1000610. [CrossRef Medline](#)
- Riecke L, Formisano E, Herrmann CS, Sack AT (2015) 4-Hz transcranial alternating current stimulation phase modulates hearing. *Brain Stimul* 8:777–783. [CrossRef Medline](#)
- Rohenkohl G, Cravo AM, Wyart V, Nobre AC (2012) Temporal expectation improves the quality of sensory processing. *J Neurosci* 32:8424–8428. [CrossRef Medline](#)
- Rosenthal R, Rubin DB (2003) R(equivalent): a simple effect size indicator. *Psychol Methods* 8:492–496. [CrossRef Medline](#)
- Schmidt-Kassow M, Heinemann LV, Abel C, Kaiser J (2013) Auditory-motor synchronization facilitates attention allocation. *Neuroimage* 82:101–106. [CrossRef Medline](#)
- Schroeder CE, Lakatos P (2009) Low-frequency neuronal oscillations as instruments of sensory selection. *Trends Neurosci* 32:9–18. [CrossRef Medline](#)
- Schroeder CE, Wilson DA, Radman T, Scharfman H, Lakatos P (2010) Dynamics of active sensing and perceptual selection. *Curr Opin Neurobiol* 20:172–176. [CrossRef Medline](#)
- Singer W (2013) Cortical dynamics revisited. *Trends Cogn Sci* 17:616–626. [CrossRef Medline](#)
- Sporns O (2014) Contributions and challenges for network models in cognitive neuroscience. *Nat Neurosci* 17:652–660. [CrossRef Medline](#)
- Sporns O, Tononi G, Edelman GM (2000) Connectivity and complexity: The relationship between neuroanatomy and brain dynamics. *Neural Netw* 13:909–922. [CrossRef Medline](#)
- Stilp CE, Kluender KR (2010) Cochlea-scaled entropy, not consonants, vowels, or time, best predicts speech intelligibility. *Proc Natl Acad Sci U S A* 107:12387–12392. [CrossRef Medline](#)
- Su YH, Pöppel E (2012) Body movement enhances the extraction of temporal structures in auditory sequences. *Psychol Res* 76:373–382. [CrossRef Medline](#)
- Tononi G, Edelman GM (1998) Consciousness and complexity. *Science* 282:1846–1851. [CrossRef Medline](#)
- Tononi G, Sporns O, Edelman GM (1994) A measure for brain complexity, relating functional segregation and integration in the nervous system. *Proc Natl Acad Sci U S A* 91:5033–5037. [CrossRef Medline](#)
- Tononi G, Edelman GM, Sporns O (1998) Complexity and coherency: integrating information in the brain. *Trends Cogn Sci* 2:474–484. [CrossRef Medline](#)
- VanRullen R, Zoefel B, Ilhan B (2014) On the cyclic nature of perception in vision versus audition. *Philos Trans R Soc Lond B Biol Sci* 369:20130214. [CrossRef Medline](#)
- Weisz N, Wühle A, Monittola G, Demarchi G, Frey J, Popov T, Braun C (2014) Prestimulus oscillatory power and connectivity patterns predispose conscious somatosensory perception. *Proc Natl Acad Sci U S A* 111:E417–E425. [CrossRef Medline](#)
- Yamamoto Y, Hughson RL (1991) Coarse-graining spectral analysis: new method for studying heart rate variability. *J Appl Physiol* 71:1143–1150. [Medline](#)
- Zar JH (2014) Biostatistical analysis. Essex: Pearson Education.