

# Frequency modulation entrains slow neural oscillations and optimizes human listening behavior

Molly J. Henry<sup>1</sup> and Jonas Obleser<sup>1</sup>

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

Edited by Nancy J. Kopell, Boston University, Boston, MA, and approved October 16, 2012 (received for review August 2, 2012)

**The human ability to continuously track dynamic environmental stimuli, in particular speech, is proposed to profit from “entrainment” of endogenous neural oscillations, which involves phase reorganization such that “optimal” phase comes into line with temporally expected critical events, resulting in improved processing. The current experiment goes beyond previous work in this domain by addressing two thus far unanswered questions. First, how general is neural entrainment to environmental rhythms: Can neural oscillations be entrained by temporal dynamics of ongoing rhythmic stimuli without abrupt onsets? Second, does neural entrainment optimize performance of the perceptual system: Does human auditory perception benefit from neural phase reorganization? In a human electroencephalography study, listeners detected short gaps distributed uniformly with respect to the phase angle of a 3-Hz frequency-modulated stimulus. Listeners’ ability to detect gaps in the frequency-modulated sound was not uniformly distributed in time, but clustered in certain preferred phases of the modulation. Moreover, the optimal stimulus phase was individually determined by the neural delta oscillation entrained by the stimulus. Finally, delta phase predicted behavior better than stimulus phase or the event-related potential after the gap. This study demonstrates behavioral benefits of phase realignment in response to frequency-modulated auditory stimuli, overall suggesting that frequency fluctuations in natural environmental input provide a pacing signal for endogenous neural oscillations, thereby influencing perceptual processing.**

pre-stimulus phase | auditory processing | EEG | FM

Neural oscillations are associated with rhythmic fluctuations in the excitation–inhibition cycle of local neuronal populations (1–3). That is, a single neuron is not equally likely to discharge in response to stimulation at all points in time. Instead, its likelihood of responding is influenced by local extracellular and membrane potentials that, in turn, are reflected in neural oscillations. Critically, these neural oscillations can be entrained by external rhythmic sensory stimulation (3, 4) or, less naturally, by rhythmic neural stimulation using transcranial magnetic stimulation (TMS) or transcranial alternating current stimulation (TACS; refs. 5 and 6). As a result, neurons are more likely to fire at temporally expected points in time. Restated, the time of peak neural sensitivity predicts the point in time at which an upcoming stimulus will occur within a framework of continued temporal regularity. Such a rhythmic neural processing mode is adaptive given the abundance of behaviorally relevant environmental auditory stimuli that are inherently rhythmic in nature and, thus, could provide a pacing signal for neural oscillations across a range of frequency bands.

Entrainment of low-frequency oscillations involves a reorganization of phase so that the optimal, in this case most excitable, phase comes into line with temporally expected critical events in the ongoing stimulus, for example tone onsets or syllable onsets (4, 7, 8). Recordings from macaque auditory cortex (3, 9, 10), human auditory cortex (11), and human electroencephalography (EEG) recordings (12–14) confirm that delta phase reorganizes in response to rhythmic auditory stimulation and that optimal phase aligns to expected times of event onsets. The results are amplified event-related potentials (ERPs) and increased firing in response to

stimulus onsets (3, 15). Behaviorally, decreased response times (RTs) to suprathreshold targets are associated with optimal delta phase (3, 14). Moreover, empirical results from vision suggest that a number of perceptual benefits are afforded to near-threshold stimuli coinciding with optimal (mostly theta and alpha) phase, including improved detection (7, 16), and faster RTs (17) (see ref. 18 for a review).

Improved perceptual processing (i.e., increased hit rate) of near-threshold stimuli occurring in the optimal phase on an entrained neural oscillation (that is, during rhythmic stimulation) has thus far not been demonstrated in the auditory domain (but see ref. 19 for evidence on phasic modulation of miss rates). This lack of evidence is unfortunate, because the rhythmic structure of speech provides a pacing signal for the reorganization of ongoing neural oscillations over a range of frequency bands (for a review, see ref. 20). For example, important information about the content of the signal is communicated by quasi-periodicity in the gamma range corresponding to fine structure and by low-frequency fluctuations (theta and delta range) corresponding to the syllable envelope and prosodic variations, respectively (21). Accordingly, evidence for neural entrainment to the slow amplitude fluctuations in tone sequences (3, 4), speech (22–24), and natural sounds (25, 26) has been demonstrated. However, a crucial missing link—a perceptual benefit for near-threshold events stemming from optimal phase alignment of an entrained oscillation in the auditory domain—has not yet been provided.

Thus far, all evidence for neural entrainment and its electrophysiological consequences has been gathered during stimulation with periodic sound onsets, by definition coupled to amplitude changes that result in an evoked response. As such, the current study made use of simple nonspeech auditory stimuli, where periodicity was communicated by frequency modulation (FM) rather than amplitude fluctuations (i.e., onsets). Although neural oscillations have been shown to be entrained by FM (27–29), the possibility that slow FM provides a pacing signal for phase alignment has not received much scientific attention (29).

Listeners detected brief auditory targets (silent gaps) embedded in the ongoing stimulus (Fig. 1). Thus, targets were not “extraneous” events, but were wholly contained within the rhythmic carrier stimulus, in much the same way as a phonemic cue in speech is not an additional, independent event, but is contained within the rhythmic speech signal. Gap locations were distributed uniformly around the 3-Hz FM cycle of the stimulus. We reasoned that if ongoing delta-band oscillations were entrained by the 3-Hz modulation, these gaps would also be uniformly distributed with respect to the phase angle of the brain oscillation. We were thus able to examine modulation profiles of behavior and evoked

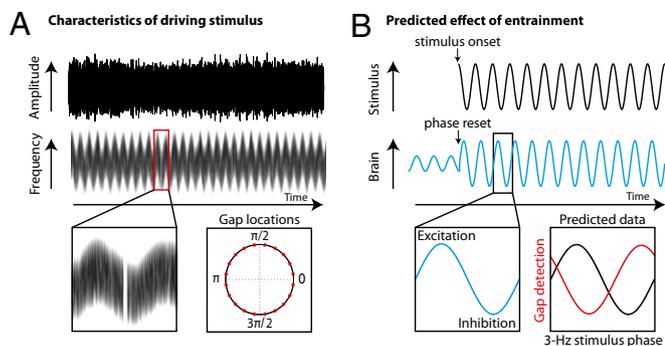
Author contributions: M.J.H. and J.O. designed research; M.J.H. performed research; M.J.H. and J.O. contributed new reagents/analytic tools; M.J.H. and J.O. analyzed data; and M.J.H. and J.O. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

<sup>1</sup>To whom correspondence may be addressed. E-mail: henry@cbs.mpg.de or obleser@cbs.mpg.de.

This article contains supporting information online at [www.pnas.org/lookup/suppl/doi:10.1073/pnas.1213390109/-DCSupplemental](http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.1213390109/-DCSupplemental).



**Fig. 1.** (A) FM stimuli for the EEG experiment. Periodicity was conveyed without fluctuations in amplitude (Top), but instead only by fluctuations in frequency (Middle). Listeners detected short silent gaps (Bottom Left) that were distributed uniformly around the 3-Hz FM cycle (Bottom Right); 2, 3, or 4 gaps were present in each 10-s stimulus. (B) Predicted neural and behavioral effects of entrainment to the FM stimulus. At stimulus onset (Top), the phase of the ongoing delta oscillation was predicted to reset to bring the neuronal oscillation into line with the driving stimulus (Middle), thereby modulating the excitation–inhibition cycle of the delta oscillation (Bottom Left). For this reason, gap detection hit rates were expected to be modulated by stimulus phase (Bottom Right).

responses by instantaneous oscillatory delta phase at the time of gap occurrence. We predicted that listening behavior (hit rates, RTs) and ERPs would be influenced by the phase of the stimulus in which the target occurred and, therefore, by the phase of the entrained delta oscillation.

## Results

Listeners detected silent gaps in 10-s complex tones that were frequency modulated at 3 Hz; a target was considered to be a “hit” if a button press occurred within 1 s of gap onset. Gaps were distributed uniformly around the 3-Hz FM cycle (in 20 possible positions); each 10-s stimulus contained two, three, or four gaps. Fig. 1A shows the frequency and amplitude envelope of an example stimulus and the 20 gap locations around the FM cycle.

We predicted that ongoing delta oscillations would be entrained by the 3-Hz frequency modulation. As a consequence, we expected that gap detection, as indexed by hit rate, would not be uniformly distributed as a function of stimulus phase, but would instead be modulated by stimulus phase (Fig. 1B). Moreover, stimulus phase effects on performance were expected to be explainable by way of the phase of the intervening neural delta (3-Hz) oscillation. We anticipated that, across listeners, optimal stimulus phase would likely be inconsistent because of individual stimulus–brain phase lags, but we expected to observe a consistent optimal delta brain phase. Finally, we also predicted that stimulus phase, by way of delta brain phase, would modulate the ERP elicited by the gap.

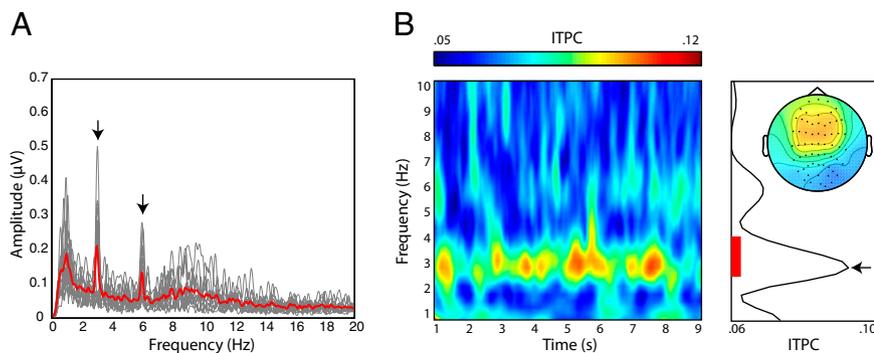
**Frequency Modulation Entrained Ongoing Delta Oscillations.** Fig. 2A shows amplitude spectra estimated from Fourier analyses for individual listeners (gray) and averaged over listeners (red), averaged over all electrodes. If ongoing neural oscillations were entrained by the 3-Hz spectral modulation, increased amplitude should be observed for the stimulation frequency, 3 Hz (and the harmonic, 6 Hz, as has been reported in response to FM stimuli; ref. 30). Therefore, two paired-samples *t* tests were conducted to test the amplitude at the two target frequencies against the average amplitude of the adjacent frequency bins (eight bins on either side of the target frequency, as recommended in ref. 30). This analysis indicated significant amplitude peaks at both 3 Hz,  $t_{(11)} = 3.32$ ,  $P = 0.008$ , and 6 Hz,  $t_{(11)} = 4.30$ ,  $P = 0.001$ .

As a main indicator of neural entrainment, intertrial phase coherence (ITPC; ref. 31) was calculated, assessing the consistency of the brain response over trials and, therefore, in response to the stimulus (Fig. 2B, averaged over listeners and electrodes). The significance of peristimulus ITPC was tested against baseline ITPC (estimated for each channel in each frequency bin and averaged over the time period from  $-1$  s to  $-0.5$  s before stimulus onset) by using a permutation *t* test (repeated measures) with a cluster-based multiple-comparisons correction (32). A single significant cluster was observed for ITPC in the 3-Hz frequency band (Fig. 2B). The topography for the significant 3-Hz cluster is in full accordance with auditory generators (33–35).

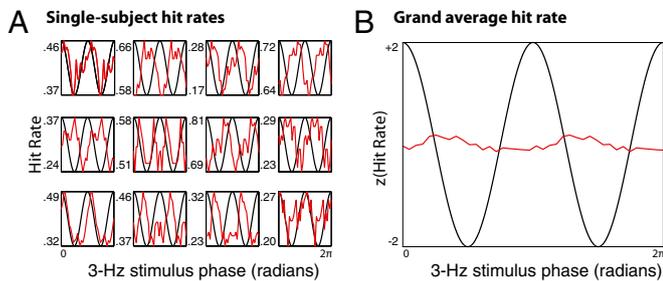
Critically, the FM phase of the stimulus at onset was randomized from trial to trial, and the reported evidence for neural entrainment at the stimulus frequency was only observable when the neural signals were realigned such that per-trial stimulus phases were consistent (SI Results and Fig. S1). Thus, phase reorganization was not simply evoked by the stimulus onset but was maintained by exact stimulus phase. In the next section, the perceptual and neural consequences of this entrainment for target detection are reported.

**Neural Entrainment Modulated Behavioral and Electrophysiological Responses to Gaps. Behavioral effects of stimulus phase.** Fig. 3 shows smoothed proportions of detected gaps (hit rate) for each of the 20 FM-phase bins for individual listeners (Fig. 3A) and averaged over listeners (Fig. 3B), superimposed on a schematic of the FM stimulus. See Fig. S2 for corroborating RT data.

Hit rates were significantly correlated with stimulus phase [average (root mean square; rms)  $\rho = 0.79$ ,  $t_{(11)} = 8.92$ ,  $P < 0.001$ ; see SI Results for details of data analysis]. This finding is simple behavioral evidence that human listening performance was, at least indirectly, modulated by neural oscillations, under the assumption confirmed above that neural oscillations were entrained by spectral stimulus fluctuations. However, it was important to rule out a trivial acoustic explanation by assessing the consistency of the phase relationship of hit rate to the stimulus across listeners. An acoustic explanation for the current results would suggest that gaps would have been inherently easier to detect when they occurred in, for example, the peak of the 3-Hz FM stimulus. We tested this



**Fig. 2.** (A) Amplitude spectrum from fast Fourier transform (FFT) of time-domain EEG signal. Amplitude in the 3-Hz and 6-Hz frequency bins was significantly larger than in the neighboring bins. Red solid line indicates the group average spectrum, gray lines show single participants’ spectra, averaged over all electrodes. (B) ITPC shown over time (Left), and averaged over time (Right), again averaged over all electrodes. The red bar indicates the frequency region in which phase coherence was significantly greater than baseline. Inset shows the topography for the significant frequency region, averaged over time; the color scale is the same as for the time-frequency representation of ITPC.



**Fig. 3.** (A) Hit rates (red) as a function of 3-Hz stimulus phase (black) for each individual listener. Two cycles of both stimulus and data have been concatenated for illustration purposes. Circular-linear correlations between stimulus phase and hit rates were significant across listeners ( $P < 0.001$ ). (B) Grand average of the z-transformed individual data in A. Across listeners, there was no consistent relation between gap detection (grand average hit rate, red) and the stimulus phase (black), ruling out acoustic explanations for the observed effect.

hypothesis formally by fitting a single-cycle cosine function to each listener's smoothed data (*SI Results*); from this function, we estimated the stimulus phase angle corresponding to peak performance. The distribution of individual "best" stimulus phases was not significantly different from a uniform distribution (Rayleigh  $z = 1.31$ ,  $P = 0.29$ ). Thus, stimulus acoustics alone are insufficient to explain the current results. This observation suggests that listeners differ in the phase lag with which their behavior related to the entraining stimulus. Next, we demonstrate that the source of interindividual variability in this lag is the intervening neural oscillation. **Behavioral effects of neural delta phase.** Fig. 4 shows smoothed hit rates from all trials of all listeners (Fig. 4C), sorted by 3-Hz delta phase (Fig. 4A) at the time of gap occurrence (0 ms); phase estimates were taken from Cz based on the central topography of the significant 3-Hz ITPC cluster. Neural delta phase was significantly correlated with hit rate [rms  $\rho = 0.76$ ,  $t_{(11)} = 8.25$ ,  $P < 0.001$ ; for full analysis details, see *SI Results*]. Hit rates were related to stimulus phase and to neural delta phase with similar strength ( $t_{(11)} = -0.49$ ,  $P = 0.63$ ). Thus, within listeners, both stimulus phase and delta phase correlated significantly (and similarly) with gap detection performance.

One important conjecture was that neural delta phase, but not stimulus phase, would predict hit rates across listeners. To test for consistency of phase effects across listeners, cosine functions were fitted to the smoothed, binned data and used to calculate the phase angle corresponding to peak performance. Estimates of optimal phase angle indicated significant clustering of optimal neural phase across listeners (Rayleigh  $z = 3.72$ ,  $P = 0.02$ ; Fig. 4G).

**Event-related potentials as a function of neural delta phase.** Fig. 4D shows an "ERP image" (i.e., all single trials from all listeners), sorted according to per-trial neural delta phase at gap onset. Similar to the results for hit rate, ERPs were systematically related to neural delta phase across listeners, but not to stimulus phase (Fig. S3). To assess phase effects on ERP components, amplitudes were extracted from the canonical time windows of the N1 (0.05–0.15 s after gap onset) and P2 (0.15–0.25 s). We were especially interested in N1 effects, because they most faithfully represent early perceptual processing of simple, to-be-detected auditory targets (36).

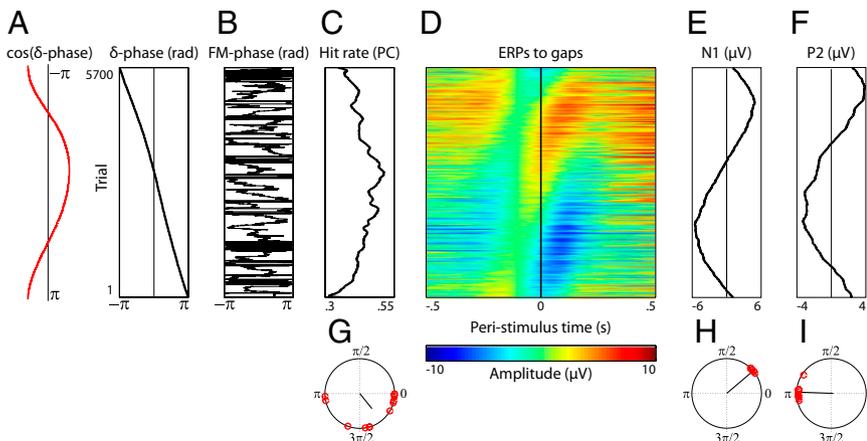
Phase effects were assessed in the same way as for hit rate data. For the N1 (Fig. 4E), the mean correlation with brain phase was extremely strong [rms  $\rho = 0.99$ ,  $t_{(11)} = 212.20$ ,  $P < 0.001$ ]. The correlation between N1 and stimulus phase (Fig. S2E) also reached significance [rms  $\rho = 0.70$ ,  $t_{(11)} = 4.57$ ,  $P < 0.001$ ], but N1 amplitude was more strongly correlated with delta phase than stimulus phase [ $t_{(11)} = 6.53$ ,  $P < 0.001$ ]. Similar results were obtained for the P2: Amplitude was significantly correlated with both brain phase [rms  $\rho = 0.98$ ,  $t_{(11)} = 32.17$ ,  $P < 0.001$ ; Fig. 4I] and stimulus phase [rms  $\rho = 0.88$ ,  $t_{(11)} = 21.19$ ,  $P < 0.001$ ; Fig. S2I], but more strongly with brain phase than with stimulus phase [ $t_{(11)} = 6.13$ ,  $P < 0.001$ ].

Optimal phase for N1 and P2 amplitude per listener was then estimated from the best-fit cosine function as a function of phase. Separate Rayleigh tests indicated that, for both N1 and P2 amplitudes, peak brain phase was consistent across listeners (N1:  $z = 11.96$ ,  $P < 0.001$ ; P2:  $z = 11.62$ ,  $P < 0.001$ ). However, peak stimulus phase was not (N1:  $z = 1.83$ ,  $P = 0.16$ ; P2:  $z = 0.26$ ,  $P = 0.78$ ).

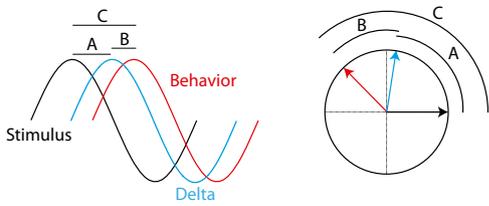
Not surprisingly, evoked responses also depended on whether the gap that elicited them would be detected. A full analysis of ERPs to hits and misses is reported in *SI Results* (and see Figs. S4 and S5). Overall, although ERPs were stronger to hits than to misses, the delta phase in which the gap was presented was a stronger predictor of the evoked signature than the behavioral response to the gap [N1:  $t_{(11)} = 8.35$ , P2:  $t_{(11)} = 10.08$ ,  $P < 0.001$ ].

**Individual Differences in Optimal Stimulus Phase Are Predictable from Intervening Brain Phase.** The relation of individual behavior to the stimulus (i.e., the stimulus–behavior relation) was inconsistent across listeners, but the relation of individual behavior to the brain (the brain–behavior relation) was consistent. We suggest that the critical variable is the intervening neural oscillation, namely the phase lag of the brain with respect to the stimulus (the stimulus–brain relation). Therefore, we expected that we could predict the stimulus–behavior relation if we knew something about individual stimulus–brain relations (Fig. 5).

This analysis required combining and comparing three lag values estimated for each listener. First, the lag of each individual's delta



**Fig. 4.** Single trials were sorted according to the instantaneous phase of the neural delta oscillation at electrode Cz (A) at the time of gap occurrence (from  $-\pi$  to  $\pi$ ). Sorting single-trial stimulus phase (B) by applying the same permutation vector revealed that stimulus phase was not consistently related to neural delta phase across listeners. However, hit rate (C) and ERPs (D–F) were. Hit rate (C), N1 amplitude (E), and P2 amplitude (F) were significantly correlated with neural delta phase ( $P < 0.001$ ). Moreover, optimal neural delta phase for hit rate (G), N1 amplitude (H), and P2 amplitude (I) was consistent across listeners ( $P < 0.02$ ). Note: this figure shows all trials for all listeners (fixed-effects), whereas all statistics reported took into account the between-subjects variance (random effects).



**Fig. 5.** A schematic depicting the reconstruction of individual stimulus-behavior lags (C) from stimulus-brain lags (A) and brain-behavior lags (B). For each listener, stimulus-brain lags were estimated from a cross-correlation analysis, whereas stimulus-behavior and brain-behavior relations were taken from analyses estimating optimal phase from hit rate data. Then, brain-behavior lags and stimulus-brain lags in radians were summed, and this sum was correlated with stimulus-behavior lags ( $P = 0.03$ ). This relationship confirmed that variability in stimulus-behavior relations is well explained by the intervening brain oscillation phase lag.

oscillation with respect to the stimulus was estimated from a cross-correlation calculated between the trial-average time-domain 3-Hz brain oscillation (bandpass-filtered between 2 Hz and 4 Hz) and the stimulus (36); we note here that a cross-correlation was chosen because here we aimed for an estimate of the lag of the brain with respect to the stimulus, rather than the consistency of the neural response. The lag corresponding to the peak correlation (converted to radians) was taken as the stimulus-brain relation (denoted “A” in Fig. 5). Second, the stimulus-behavior relation was taken as the estimated lag parameter from the cosine fits to hit rates as a function of stimulus phase (“C” in Fig. 5). Third, the brain-behavior relation was similarly estimated as the lag of cosine-fitted hit rates as a function of delta phase (“B” in Fig. 5).

We summed the phase lags for the stimulus-brain (A) and brain-behavior (B) relations. We assumed that the combination of these two phase lags should predict the stimulus-behavior lag (C). That is, we assumed we could account for individual variability in the stimulus-behavior lag by taking into account the intervening brain signal. Accordingly, we correlated the summed stimulus-brain and brain-behavior lags (A+B) with the stimulus-behavior lag (C) by using a circular-circular correlation and found that the stimulus-behavior lag (C) was predictable from the intervening brain relation between the stimulus and behavior [ $\rho_{(22)} = 0.46$ ,  $P = 0.033$ , one-tailed].

## Discussion

The current study demonstrated that low-frequency (3-Hz) delta oscillations are entrained by spectral fluctuations in nonspeech auditory stimuli. In turn, instantaneous phase of both the stimulus and the entrained neural oscillation determined optimal listening behavior (here: gap detection performance). Moreover, entrained delta oscillations shaped the auditory potential evoked by the gap (ERP), and delta phase was more strongly predictive of ERP amplitude than whether the gap was detected. We interpret the current results as reflecting low-frequency oscillation of a neural excitation-inhibition cycle, which governs listening performance in a difficult (near-threshold) auditory perception task.

Critically, optimal stimulus phase varied across listeners, ruling out a trivial acoustic explanation of the current results. Instead, optimal neural delta phase was consistent across listeners. Moreover, we were able to explain individual differences in stimulus phase effects on behavior by taking into account variation across listeners in the relation of the entrained delta oscillation to the driving stimulus.

Perceptual benefits (enhanced detection; faster RTs) have been shown for near-threshold visual stimuli occurring in the optimal phase of theta or alpha oscillations (16, 17, 37). Moreover, in the auditory domain, delta phase reorganization has been demonstrated in response to rhythmic tone sequences, resulting in a more

vigorous neural response to tone onsets (3, 14) and modulation of RTs to suprathreshold stimuli (14). A recent study demonstrated modulation of miss rates by low frequency (2–6 Hz) oscillatory phase for near-threshold auditory targets embedded in an ongoing aperiodic stimulus (19).

However, the present data demonstrate that phase reorganization in response to a periodic auditory stimulus influences perceptual processing of near-threshold target stimuli in a periodic fashion. Moreover, although phase locking to FM stimuli has been observed (27–29), this study shows optimal-phase effects due to phase locking to an auditory stimulus modulated only spectrally (i.e., without any rhythmic amplitude fluctuations). Thus, the current results are suggestive that spectral fluctuations in natural auditory stimuli may act as a pacing signal for low-frequency oscillations and may, in turn, influence auditory perception.

**Entrainment in the Absence of Envelope Information.** Much emphasis is placed on amplitude fluctuations as the time-varying signal by which ongoing neural oscillations might be entrained (22, 23). This strong emphasis is epitomized by a recently proposed model in which speech comprehension is formalized in terms of the goodness of entrainment of ongoing theta-band oscillations to the amplitude (syllable) envelope of speech (38). In this model, the active adjustment of the stimulus-brain relation is proposed to be supported by a phase-reset mechanism, which critically relies on acoustic onsets to evoke the reset response (ref. 20 for further discussion of this point). For this reason, envelope cues (at the syllabic rate) are proposed to be of utmost importance for entrainment of ongoing brain oscillations to an acoustic signal.

Here, we chose to make use of a FM stimulus for three reasons. First, although it is commonly acknowledged that speech contains slow fluctuations in the frequency domain, the possibility that these fluctuations might provide a pacing signal for phase reorganization of low-frequency oscillations has not received much scientific attention. This study demonstrates behavioral consequences of delta phase reorganization in response to periodic spectral changes. It is noteworthy in this regard that we did not see evidence for phase reorganization when we did not realign trials with respect to stimulus phase.

Second, the constant amplitude over the cycle of the FM allowed us to uniformly distribute targets around the stimulus phase, and thereby construct a modulation profile for hit rates and ERPs as a function of entrained delta phase. Because the phase of the entrained neural delta oscillation was predictable from the phase of the FM stimulus, we were capable of targeting specific neural delta phases to assess modulatory effects on perception. This technique is similar to, but less obtrusive than, paradigms that drive brain oscillations with TACS or TMS (5, 6) and present targets at known phases of the entrained oscillation.

Third, an advantage of using an entraining stimulus without amplitude fluctuations (i.e., onsets) is the avoidance of complications that come with attempting to measure “proper entrainment” separately from periodic evoked responses to rhythmically presented stimuli. It has been argued that “entrainment” may reflect no more than rhythmic evoked responses (39). When the temporal context is defined based on discrete events (e.g., tones), these two possibilities are not separable. This confound arises because individual ERPs occur in response to event onsets at precisely the stimulation frequency, leading to increased spectral power and phase locking at the stimulation frequency in the (time-)frequency domain. In the present study, rhythmic information was communicated by frequency modulation. Thus, we argue that increased 3-Hz power and phase coherence did not result from periodic evoked responses but rather reflect entrainment of neural oscillations.

At the level of an individual frequency-tuned neuronal population, frequency modulation constitutes a local “onset” and, thus, an evoked response, which follows from the tonotopic organization

of the auditory system, beginning with cochlear output. An auditory stimulus with continuously varying frequency serially stimulates neighboring cortical locations (40, 41); therefore, a single frequency-tuned neuronal population would respond periodically to an FM stimulus, with increased firing coinciding with the time point at which the frequency trajectory intersects the characteristic frequency of that cortical location. From this perspective, it could be argued that on a very local spatial scale, an FM stimulus is not so different from a periodic tone sequence, in terms of recurrent onsets and offsets.

However, our behavioral results would be very difficult to explain based on local excitation by frequency modulation. Because the pattern of evoked excitations is continuous—that is, based on continuous stimulation—an optimal stimulus phase for gap detection would imply a privileged set of frequency-tuned neurons that respond best to near-threshold auditory targets. Moreover, because optimal stimulus phase varied across listeners, the privileged neuronal population would differ between listeners in terms of frequency tuning.

Overall, we suggest that a distinction can be made between local excitation evoked by transient stimulation of a frequency-tuned neuronal population (FM) and global excitation evoked by sound onsets and associated with a perceptual onset (AM). However, future research on FM-induced entrainment and phase reorganization will profit by using intracranial recording techniques in human or animal subjects.

**Interpreting Optimal Phase.** In the current study, we found that detection performance and N1 amplitudes peaked for gaps that were presented in the rising phase, near the peak, of the 3-Hz delta oscillation (Fig. 4). These results are perhaps unexpected based on the idea that local neuronal populations should be most excitable when the slow oscillation indexing local potentials is at its most negative point (i.e., in the trough of the oscillation), as has been demonstrated recording directly from macaque auditory cortex (A1; 3, 4, 9). In contrast, the current study relies on EEG, in which each single electrode measures summed electric potential fluctuations through the skull. Thus, the EEG signal is comparably impure. Moreover, the absolute phase value of an EEG signal depends on the choice of reference (42). Thus, we exercise some caution when interpreting absolute values of optimal phase.

With this cautionary note in mind, one of the most interesting findings emerging from the current study was based on relative phase. That is, the stimulus phase in which individual listeners performed best was variable (Fig. 3). However, we were able to predict peak stimulus phase (i.e., the stimulus–behavior lag) from a simple combination of the stimulus–brain lag and the brain–behavior lag. This result confirms our hypothesis that the entrained delta oscillation was the intervening variable between the stimulus and the pattern of behavioral results observed for each listener.

**Neural Oscillations and Temporal Context.** The role of the rhythmic context in which an event is situated recently has gained much attention, and a theoretical proposal regarding rhythmic attention captures well this zeitgeist (8). The critical idea behind this proposal is that attention can behave in either a “rhythmic” or a “continuous” mode, with the former being the default mode of operation. Under the rhythmic processing scheme, neuronal sensitivity fluctuates and is heightened at times when stimuli are most likely to occur. The rhythmic mode can be contrasted with an energetically more expensive, constant state of high neuronal sensitivity that subserves continuous monitoring for a temporally unpredictable stimulus. Because of the imbalance in energy demands, it is suggested that the rhythmic mode is preferred. The preference for a rhythmic processing mode fits well with the observation that many natural auditory signals are, to varying extents, rhythmic. Moreover, it can be noted that many sensory behaviors actively push perception into a rhythmic mode (43). For example, in vision, saccades occur rhythmically with a rate near 3 Hz; visual

input coinciding with stable viewing is enhanced, whereas visual input during a saccade is suppressed, consistent with rhythmic attending. Moreover, in nonhuman animals, respiration tied to sniffing (olfaction) and whisking behavior (somatosensation) are both rhythmic, with periods of excitation interleaved with periods of inhibition.

This proposal fits well with a theory of rhythmic attending developed by Jones and colleagues (44–47), referred to as dynamic attending theory. In brief, dynamic attending theory proposes that attention waxes and wanes rhythmically and that events coinciding with attentional “peaks” are privy to more thorough processing. Moreover, attentional rhythms are entrained by environmental rhythms. Thus, better processing is afforded to stimuli occurring at points in time that are expected based on the rhythmic context in which they are presented. Indeed, perception of onset time (47), pitch perception (48), and gap detection (49) are better for temporally expected relative to unexpected stimuli. Although the proposed mechanism underlying these perceptual benefits is oscillation of an “attentional pulse” as opposed to a neural oscillation, these ideas lay important groundwork for the burgeoning interest in perceptual consequences of neural entrainment by environmental rhythms.

## Conclusions

In conclusion, this study demonstrates that successful detection of (and response times and ERPs to) difficult-to-detect auditory stimuli depends on the instantaneous phase of neural delta oscillations that are entrained by a spectrally modulated stimulus. This observation constitutes an important generalization of the neural entrainment hypothesis to rhythmic stimuli in the absence of salient stimulus onsets and offsets. We suggest that low-frequency (i.e., delta-range) fluctuations of natural auditory stimuli in the frequency domain, for example prosodic or melodic contour information, act as a pacing signal by which slow neural delta oscillations are entrained, thereby optimizing human listening behavior.

## Experimental Procedures

**Participants.** Twelve normal-hearing, native German speakers (age 21–32, 6 female) took part in the study. Data from another four participants were collected but were discarded because of a high degree of noise in the EEG recording ( $n = 3$ ) or inability to perform the gap detection task ( $n = 1$ ). Participants received financial compensation of 15 €. Written informed consent was obtained from all participants. The procedure was approved 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 at a sampling rate of 60,000 Hz. Stimuli were 10-s complex tones frequency-modulated at a rate of 3 Hz and a depth of 37.5% (Fig. 1). Complex carrier signals were centered on one of three frequencies (800, 1,000, 1,200 Hz) and comprised of 30 components sampled from a uniform distribution with a 500-Hz range. The amplitude of each component was scaled linearly based on its inverse distance from the center frequency; that is, the center frequency itself was the highest-amplitude component, and component amplitudes decreased with increasing distance from the center frequency. The onset phase of the stimulus was randomized from trial to trial, taking on one of eight values ( $0, \pi/4, \pi/2, 3\pi/4, \pi, 5\pi/4, 3\pi/2, 7\pi/4$ ). All stimuli were rms amplitude normalized.

Two, three, or four silent gaps were inserted into each 10-s stimulus (gap onset and offset were gated with 3-ms half-cosine ramps) without changing the duration of the stimulus. Each gap was chosen to be centered in 1 of 20 equally spaced phase bins into which each single cycle of the frequency modulation was divided. Gaps were placed in the final 9 s of the stimulus, with the constraint that two gaps could not occur within 667 ms (i.e., 2 cycles) of each other.

**Procedure.** Gap duration was first titrated for each individual listener; individual thresholds ranged between 10 ms and 18 ms. For the main experiment, EEG was recorded while listeners detected gaps embedded in 10-s-long FM stimuli. Each stimulus contained 2, 3, or 4 gaps, and listeners responded with a button-press when they detected a gap. Listeners were instructed to respond as quickly as possible when they detected a gap, and a response was considered to be a “hit” when it occurred within 1 s of a gap. Overall, each listener heard 70 stimuli per carrier frequency, for a total of

210 trials. For each of the 20 FM-phase bins, 30 gaps were presented, 10 per carrier frequency, for a total of 600 targets. On average, the experiment lasted between 90 and 120 min including preparation of the EEG. See [SI Experimental Procedures](#) for additional procedural information.

**Data Acquisition and Analysis. Behavioral data.** Behavioral data were recorded online by Presentation software (Neurobehavioral Systems). Hits were defined as button-press responses that occurred no more than 1 s after the occurrence of a gap. Hit rates and RTs were calculated separately for each of the 20 FM-phase bins. See [SI Experimental Procedures](#) for additional details. **Electroencephalogram data.** The EEG was recorded from 64 Ag–AgCl electrodes mounted on a custom-made cap (Electro-Cap International), according to the modified and expanded 10–20 system. Signals were recorded continuously with a passband of DC to 200 Hz and digitized at a sampling rate of 500 Hz. The reference electrode was the left mastoid. Bipolar horizontal and vertical EOGs were recorded for artifact rejection purposes. Electrode resistance was kept under 5 kΩ.

All EEG data were analyzed offline by using Fieldtrip software ([www.ru.nl/fcdonders/fieldtrip](http://www.ru.nl/fcdonders/fieldtrip); ref. 50), and custom Matlab (Mathworks) scripts. After

artifact rejection, full stimulus epochs were analyzed in the frequency and time–frequency domains to examine oscillatory brain responses entrained by the 3-Hz stimulation. Specifically, we examined amplitude spectra and inter-trial phase coherence (ITPC).

For target epochs, ERPs were examined in the time domain, and power, ITPC, and a bifurcation index (37) were analyzed in the time–frequency domain. For each of these dependent measures, responses to detected (hits) and undetected (misses) targets were compared directly. Moreover, single-trial complex output from the Wavelet convolution was also used to estimate the phase angle in the 3-Hz band at the time of gap occurrence, then single-trial ERPs were sorted with respect to the delta phase angle to characterize delta phase effects. For hit rates, RTs, N1 amplitudes, and P2 amplitudes, the optimal delta or stimulus phase angle was estimated based on a single-cycle cosine fit to the data. For additional data analysis information, please see [SI Experimental Procedures](#).

**ACKNOWLEDGMENTS.** We thank Heike Boethel for help with the data acquisition. This work was supported by the Max Planck Society, Germany, through a Max Planck Research Group grant (to J.O.).

- Bishop GH (1933) Cyclic changes in the excitability of the optic pathway of the rabbit. *Am J Physiol* 103:213–224.
- Buzsáki G, Draguhn A (2004) Neuronal oscillations in cortical networks. *Science* 304(5679):1926–1929.
- Lakatos P, et al. (2005) An oscillatory hierarchy controlling neuronal excitability and stimulus processing in the auditory cortex. *J Neurophysiol* 94(3):1904–1911.
- Lakatos P, Karmos G, Mehta AD, Ulbert I, Schroeder CE (2008) Entrainment of neuronal oscillations as a mechanism of attentional selection. *Science* 320(5872):110–113.
- Thut G, et al. (2011) Rhythmic TMS causes local entrainment of natural oscillatory signatures. *Curr Biol* 21(14):1176–1185.
- Zaehle T, Rach S, Herrmann CS (2010) Transcranial alternating current stimulation enhances individual alpha activity in human EEG. *PLoS ONE* 5(11):e13766.
- Mathewson KE, Gratton G, Fabiani M, Beck DM, Ro T (2009) To see or not to see: Prestimulus alpha phase predicts visual awareness. *J Neurosci* 29(9):2725–2732.
- Schroeder CE, Lakatos P (2009) Low-frequency neuronal oscillations as instruments of sensory selection. *Trends Neurosci* 32(1):9–18.
- Lakatos P, Chen C-M, O’Connell MN, Mills A, Schroeder CE (2007) Neuronal oscillations and multisensory interactions in primary auditory cortex. *Neuron* 53(2):272–292.
- Lakatos P, et al. (2009) The leading sense: Supramodal control of neurophysiological context by attention. *Neuron* 64(3):419–430.
- Besle J, et al. (2011) Tuning of the human neocortex to the temporal dynamics of attended events. *J Neurosci* 31(9):3176–3185.
- Başar E, Başar-Eroglu C, Rosen B, Schütt A (1984) A new approach to endogenous event-related potentials in man: Relation between EEG and P300-wave. *Int J Neurosci* 24(1):1–21.
- Başar E, Stampfer HG (1985) Important associations among EEG-dynamics, event-related potentials, short-term memory and learning. *Int J Neurosci* 26(3-4):161–180.
- Stefanics G, et al. (2010) Phase entrainment of human delta oscillations can mediate the effects of expectation on reaction speed. *J Neurosci* 30(41):13578–13585.
- Haig AR, Gordon E (1998) EEG alpha phase at stimulus onset significantly affects the amplitude of the P3 ERP component. *Int J Neurosci* 93(1-2):101–115.
- Busch NA, VanRullen R (2010) Spontaneous EEG oscillations reveal periodic sampling of visual attention. *Proc Natl Acad Sci USA* 107(37):16048–16053.
- Drewes J, VanRullen R (2011) This is the rhythm of your eyes: The phase of ongoing electroencephalogram oscillations modulates saccadic reaction time. *J Neurosci* 31(12):4698–4708.
- VanRullen R, Busch NA, Drewes J, Dubno JR (2011) Ongoing EEG phase as a trial-by-trial predictor of perceptual and attentional variability. *Frontiers in Psychology* 2:1–9.
- Ng BSW, Schroeder T, Kayser C (2012) A precluding but not ensuring role of entrained low-frequency oscillations for auditory perception. *J Neurosci* 32(35):12268–12276.
- Schroeder CE, Lakatos P, Kajikawa Y, Partan S, Puce A (2008) Neuronal oscillations and visual amplification of speech. *Trends Cogn Sci* 12(3):106–113.
- Rosen S (1992) Temporal information in speech: Acoustic, auditory and linguistic aspects. *Philos Trans R Soc Lond B Biol Sci* 336(1278):367–373.
- Ahissar E, et al. (2001) Speech comprehension is correlated with temporal response patterns recorded from auditory cortex. *Proc Natl Acad Sci USA* 98(23):13367–13372.
- Luo H, Poeppel D (2007) Phase patterns of neuronal responses reliably discriminate speech in human auditory cortex. *Neuron* 54(6):1001–1010.
- Peelle JE, Gross J, Davis MH (May 17, 2012) Phase-locked responses to speech in human auditory cortex are enhanced during comprehension. *Cereb Cortex*, 10.1093/cercor/bhs118.
- Kayser C, Montemurro MA, Logothetis NK, Panzeri S (2009) Spike-phase coding boosts and stabilizes information carried by spatial and temporal spike patterns. *Neuron* 61(4):597–608.
- Ng BSW, Logothetis NK, Kayser C (February 17, 2012) EEG phase patterns reflect the selectivity of neural firing. *Cereb Cortex*, 10.1093/cercor/bhs031.
- Luo H, Wang Y, Poeppel D, Simon JZ (2006) Concurrent encoding of frequency and amplitude modulation in human auditory cortex: MEG evidence. *J Neurophysiol* 96(5):2712–2723.
- Millman RE, Prendergast G, Kitterick PT, Woods WP, Green GGR (2010) Spatiotemporal reconstruction of the auditory steady-state response to frequency modulation using magnetoencephalography. *Neuroimage* 49(1):745–758.
- Obleser J, Herrmann B, Henry MJ (2012) Neural oscillations in speech: Don’t be enslaved by the envelope. *Front Human Neurosci* 6:250.
- Picton TW, John MS, Dimitrijevic A, Purcell D (2003) Human auditory steady-state responses. *Int J Audiol* 42(4):177–219.
- Lachaux J-P, Rodriguez E, Martinerie J, Varela FJ (1999) Measuring phase synchrony in brain signals. *Hum Brain Mapp* 8(4):194–208.
- Maris E, Oostenveld R (2007) Nonparametric statistical testing of EEG- and MEG-data. *J Neurosci Methods* 164(1):177–190.
- Budd TW, Barry RJ, Gordon E, Rennie PT (1998) Decrement of the N1 auditory event-related potential with stimulus repetition: Habituation vs. refractoriness. *Int J Psychophysiol* 31(1):51–68.
- Lenz D, Schadow J, Thaeig S, Busch NA, Herrmann CS (2007) What’s that sound? Matches with auditory long-term memory induce gamma activity in human EEG. *Int J Psychophysiol* 64(1):31–38.
- Oades RD, Zerbin D, Dittmann-Balcar A (1995) The topography of event-related potentials in passive and active conditions of a 3-tone auditory oddball test. *Int J Neurosci* 81(3-4):249–264.
- Skoe E, Kraus N (2010) Auditory brain stem response to complex sounds: A tutorial. *Ear Hear* 31(3):302–324.
- Busch NA, Dubois J, VanRullen R (2009) The phase of ongoing EEG oscillations predicts visual perception. *J Neurosci* 29(24):7869–7876.
- Giraud A-L, Poeppel D (2012) Cortical oscillations and speech processing: Emerging computational principles and operations. *Nat Neurosci* 15(4):511–517.
- Capilla A, Pazo-Alvarez P, Darriba A, Campo P, Gross J (2011) Steady-state visual evoked potentials can be explained by temporal superposition of transient event-related responses. *PLoS ONE* 6(1):e14543.
- Horikawa J, Nasu M, Taniguchi I (1998) Optical recording of responses to frequency-modulated sounds in the auditory cortex. *Neuroreport* 9(5):799–802.
- Weisz N, Wienbruch C, Hoffmeister S, Elbert T (2004) Tonotopic organization of the human auditory cortex probed with frequency-modulated tones. *Hear Res* 191(1-2):49–58.
- Núñez PL (2005) *Electric Fields in the Brain: The Neurophysics of EEG* (Oxford Univ Press, Oxford).
- Schroeder CE, Wilson DA, Radman T, Scharfman H, Lakatos P (2010) Dynamics of active sensing and perceptual selection. *Curr Opin Neurobiol* 20(2):172–176.
- Jones MR (1976) Time, our lost dimension: Toward a new theory of perception, attention, and memory. *Psychol Rev* 83(5):323–355.
- Jones MR (2004) Attention and timing. *Ecological Psychoacoustics*, ed Neuhoff JG (Elsevier, San Diego), pp 49–85.
- Large EW, Jones MR (1999) The dynamics of attending: How people track time-varying events. *Psychol Rev* 106(1):119–159.
- 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(6):1102–1125.
- Jones MR, Moynihan H, MacKenzie N, Puenté J (2002) Temporal aspects of stimulus-driven attending in dynamic arrays. *Psychol Sci* 13(4):313–319.
- Lange K (2009) Brain correlates of early auditory processing are attenuated by expectations for time and pitch. *Brain Cogn* 69(1):127–137.
- Oostenveld R, Fries P, Maris E, Schoffelen J-M (2011) FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Comput Intell Neurosci* 2011:1–9.

# Supporting Information

Henry and Obleser 10.1073/pnas.1213390109

## SI Results

**Electrophysiological Correlates of Gap Detection.** After demonstrating the dependence of gap detection on delta brain phase, we examined the electrophysiological correlates of detected versus undetected gaps. Fig. S4A shows event-related potentials (ERPs) elicited by detected and undetected gaps, averaged over all electrodes, with regions of significant difference marked in red [based on the Fieldtrip-implemented, false discovery rate (FDR)-corrected, paired-samples permutation  $t$  test, with cluster correction; ref. 1]. It is clear that ERPs to detected gaps were overall larger than ERPs to undetected gaps.

Results for power changes (Fig. S4B) and intertrial phase coherence (ITPC) (Fig. S4C) in response to gaps revealed converging results. That is, both metrics revealed significant differences spanning the delta, theta, and alpha bands (2–12 Hz) in the time window corresponding to the gap-evoked response. With respect to power, a significant enhancement was observed in a single cluster spanning the delta and theta bands (2–8 Hz). Suppression was observed in the alpha (two clusters, one impinging on theta: 9–11 Hz, 6–10 Hz) and beta (one cluster: 15–25 Hz) bands, that was stronger for detected than for undetected gaps. Similarly, ITPC in the delta, theta, and alpha bands (single significant cluster: 2–15 Hz frequency range) was increased for detected relative to undetected gaps, also in the time range of the ERP. Thus, both power and phase coherence results suggest enhancement of the phase-locked evoked response to detected relative to undetected gaps.

We also calculated a bifurcation index (2), which indexes the consistency of the phase reset due to detected versus undetected gaps (Fig. S4D). Negative values indicate stronger phase concentration for one target type (e.g., hits) than for the other (e.g., misses), whereas positive values indicate phase consistency for both hits and misses, but with different preferred phase angles. The value of the bifurcation index was significantly negative relative to the pretarget baseline period (–1 to –0.5 s) in the delta, theta, and alpha frequency bands (2–12 Hz) for 600 ms after target occurrence. Taken together with the ITPC results, gap-evoked responses were consistent following detected targets, whereas undetected targets did not reset phase in the 2- to 12-Hz range to a consistent angle.

**Delta Phase is a Better Predictor of ERP Magnitude than Target Detection.** We also compared the degree to which ERP magnitudes correlated with hit rates versus delta phase. As would be expected, N1 [ $t_{(11)} = 5.21$ ,  $P > 0.001$ ] and P2 [ $t_{(11)} = 1.98$ ,  $P = 0.07$ ] amplitudes were significantly correlated with hit rate (this was only marginal for P2 amplitude). We wanted to rule out the possibility that the relation between the neural delta oscillation and ERP was merely a by-product of more detected targets (with therefore larger ERPs) occurring near the peak of the delta oscillation. If modulation of ERPs by delta phase was simply due to more hits occurring in some delta phases than in others, we would have expected the hit rate–ERP correlation to be greater than or equal to the delta phase–ERP correlation. However, delta phase [N1:  $t_{(11)} = 8.35$ , P2:  $t_{(11)} = 10.08$ ] predicted ERP amplitudes better than hit rate ( $P < 0.001$ ), indicating that phase effects were not simply the result of more detected targets for some delta phases relative to others.

**Conditional Probability of Gap Detection.** We also conducted an analysis in which we calculated hit rates as a function of the time since the previous target occurrence. Each stimulus contained between two and four gaps, and the minimum duration between gaps was 667 ms. Thus, gaps occasionally occurred in the time

window of the phase-reset response evoked by the previous gap (the “ERP window”). Because hits and misses had significantly different gap-evoked response signatures, it was possible that estimating instantaneous phase in this time window may have been in part responsible for the delta phase effects we observed.

Thus, we examined detection performance for only those gaps that followed another target within the same 10-s stimulus. We binned these gaps based on the duration separating them from the preceding target (1–6 s; we ignored gaps occurring after a longer duration because there were too few of these trials to provide a meaningful comparison). Moreover, we split the gaps according to whether the preceding target was a hit (and thus elicited a phase reset) or a miss (and was less likely to elicit a phase reset). Fig. S5 shows hit rates as a function of time since the preceding gap, split according to whether the previous target was a hit or a miss. A repeated-measures ANOVA revealed a significant main effect of time since the previous gap ( $P = 0.001$ ), which was qualified by a significant interaction with whether the previous gap was a hit or miss ( $P = 0.02$ ). The interaction was driven by lower hit rates for gaps occurring within 1 s of a detected gap than an undetected gap ( $P = 0.005$ ); pairwise comparisons of hit rates for gaps following hits vs. misses did not reach significance at any other time point ( $P > 0.2$ ).

This analysis indicated that gaps occurring within the ERP time window of another detected gap were more likely to be missed than if the preceding gap was not detected and, thus, did not produce a strong phase reset. To rule out that the observed delta phase effects were an artifact of this result, we removed all trials on which the gap occurred within 1 s of another gap, regardless of whether the preceding gap was a hit or miss. Then, we replicated our analysis that involved sorting trials by delta phase (Fig. 4). Critically, we found that even without trials where gaps occurred within an ERP window, delta phase significantly predicted hit rate ( $t = 18.98$ ,  $P < 0.001$ ,  $\text{rms } \rho = 0.87$ ), N1 amplitude ( $t = 183.22$ ,  $P < 0.001$ ,  $\text{rms } \rho = 0.99$ ), and P2 amplitude ( $t = 28.16$ ,  $P < 0.001$ ,  $\rho = 0.98$ ). We thus conclude that our delta phase results were not simply due to occasionally presenting a target in the time window of the phase-reset response evoked by the previous gap.

## SI Experimental Procedures

**Procedure.** Titration of individual gap duration was accomplished by using a two-down one-up adaptive tracking procedure, which converged on the gap duration corresponding to 70.7% correct in a three-alternative forced-choice (3AFC) gap detection task (3, 4) in MATLAB on a MacBook Pro laptop (Apple). On each trial, listeners were presented with three 1-s FM stimuli, constructed according to the same constraints as the stimuli in the experiment proper, and indicated which of the three contained a gap. The gap was always temporally centered in the stimulus. The starting phase of each of the three stimuli was randomized independently and could take on any value between 0 and  $2\pi$ . Thus, gaps occurred during the thresholding procedure equally often at all stimulus phases. Thresholds were therefore an approximate average of individual thresholds corresponding to different stimulus phase locations. Listeners completed three blocks of the adaptive tracking procedure in ~15 min. Twelve reversals were completed during each block, and thresholds were taken as the arithmetic average of the final eight reversals. The final individual gap duration was taken as the average of the three estimates from the individual blocks.

For the main experiment, listeners were seated in front of a black-screen computer monitor in an EEG cabin. They registered responses with a button box, which they were permitted to hold in their

lap or set on the table in front of them. Each trial was initiated with a button press, which was followed by the appearance of a fixation cross after a variable interval (centered on 1.5 s), and after another variable interval (centered on 1.5 s) by the onset of the sound. The experiment was self-paced, in that listeners were allowed to break as long as they wished before initiating the next trial.

**Data Analysis. Behavioral data.** Behavioral performance was modulated by the FM phase at which the to-be-detected gap occurred. To confirm this observation, separate circular-linear correlations were calculated between stimulus phase and hit rate for each listener. To test the strength of these correlations across listeners, correlations were first converted to coefficients of determination ( $R^2$ ) by squaring, and then arcsine transformed to overcome nonnormality due to bounding of circular-linear correlations between 0 and 1 (5). Squared, transformed correlation coefficients were then tested against the null hypothesis of zero correlation.

**EEG data.** Two sets of epochs-of-interest were defined. Full-stimulus epochs were defined as 1.5 s preceding and 11.5 s after the sound onset to capture the full 10-s stimulus. Target epochs were defined as 2 s preceding and 2 s after each gap occurrence. Data were bandpass filtered between 0.1 Hz and 100 Hz, then artifacts were rejected in two steps. First, independent components analysis (ICA) was used to eliminate blinks, electrooculogram (EOG), and muscle activity. For full-stimulus epochs, this procedure resulted in removal of  $M = 4.27 \pm 1.37$  components (range: 2–7), and for target epochs,  $M = 4.92 \pm 1.44$  components were removed (range: 3–7). Second, individual trials were automatically rejected by using a threshold-based rejection routine with a threshold of 120  $\mu\text{V}$ . For full-stimulus epochs, after ICA,  $5.2 \pm 6.1\%$  of trials were removed (range 1–38 of 210 trials), and for target epochs,  $21 \pm 11\%$  of trials were removed (range 28–262 of 600 trials).

To examine oscillatory brain responses entrained by the 3-Hz stimulation, full-stimulus epochs were analyzed in the frequency domain. It should be noted that the starting phase of the FM stimulus was randomized from trial to trial. Therefore, before conducting frequency-domain analyses, brain responses were shifted in time so that the FM stimulus on each trial would have been perfectly phase-locked across trials. To estimate power in each frequency band, a fast Fourier transform (FFT) was performed on the trial-averaged time-domain data, after high-pass filtering at 0.5 Hz to reduce  $1/f$  noise and multiplication with a Hann window. The single-trial time-domain data were submitted to a time-frequency analysis by using the Fieldtrip-implemented version of the Wavelet approach using Morlet wavelets (6, 7), with which the time series were convolved. Wavelet-based approaches to estimating time-frequency representations of EEG data form a good compromise between frequency and time resolution. Here, wavelet size varied with frequency linearly from three to seven cycles over the range from 1 to 15 Hz (Fig. 2 shows only up to 10 Hz). The resulting complex values were used to estimate ITPC (8) for each channel, for each frequency-time bin. ITPC was calculated according to the formula

$$\text{ITPC}_{(c,f,t)} = \frac{1}{N} \left| \sum_{k=1}^N e^{i\theta(c,f,t,k)} \right|,$$

where  $\theta(c,f,t,k)$  is the single-trial, instantaneous phase angle of the ongoing oscillation on a single trial ( $k$ ). The value of phase coherence is equal to the resultant vector length of the sample of phase angles and is bounded between a minimum of 0 and a maximum of 1. Because phase coherence values are bounded and are therefore not normally distributed, values were arcsine-transformed (5) before being submitted to statistical analysis.

Target epochs were first analyzed in the time domain; data for detected and undetected gaps were time locked with respect to gap onset, then low-pass filtered below 15 Hz and subjected to a paired-samples  $t$  test with a cluster-based correction for multiple comparisons (1). Target epochs were also subjected to a wavelet analysis; the complex output of the wavelet convolution was used to estimate power and ITPC separately for detected and undetected gaps, which were again compared with a paired-samples  $t$  test with cluster correction. Finally, ITPC values were used to estimate a bifurcation index (2), which can be calculated according to the formula:

$$(\text{ITPC}_d - \text{ITPC}_T) \times (\text{ITPC}_u - \text{ITPC}_T),$$

where  $\text{ITPC}_d$  refers to ITPC across all trials on which the target was detected,  $\text{ITPC}_u$  refers to ITPC across undetected target trials, and  $\text{ITPC}_T$  refers to the total ITPC over all trials. For the bifurcation index, negative values indicate significant phase concentration of either detected or undetected target trials (but not both), whereas positive values indicate significant phase concentration for both trial types, but with a different mean phase angle. Values near zero indicate that either both trial types are phase locked with the same mean angle or phase distributions for both trial types are uniform.

Trials were low-pass filtered with a 100-sample kernel. To estimate the relationship of ERP component amplitudes to neural delta phase at the time of target occurrence, mean amplitudes were extracted from time windows centered on the canonical N1 (50–150 ms) and P2 (150–250 ms) components. Random-effects analyses involved calculating circular-linear correlations between delta phase and each of these dependent measures for each listener. The correlation coefficients were first transformed to coefficients of determination by squaring, so that they would be additive and amenable to statistics. Then, because circular-linear correlation coefficients are bounded between 0 and 1, coefficients of determination were arcsine transformed before being submitted to single-sample  $t$  tests against 0. This analysis was repeated for stimulus phase (Fig. S3).

Optimal phase was estimated for each dependent variable [hit rate, response time (RT), ERP amplitude] with respect to both the stimulus and the brain by using the following procedure. Single-trial phase values were used to sort hits and ERPs for single trials into 20 bins corresponding to the same phase values as shown in Fig. 1. N1 and P2 time windows were defined the same as above. Thus, as for stimulus phase, the result was 20 hit rates, RTs, and ERPs, corresponding to 20 phase bins.

Binned data (for both stimulus and brain phase) were smoothed by using a circular smoothing method with a five-sample kernel. Then, for each listener, a single-cycle cosine function was fit to the smoothed data by using a MATLAB-implemented least squares routine (lsqcurvfit):

$$f(j) = \cos(2\pi f_m t(j) + \phi),$$

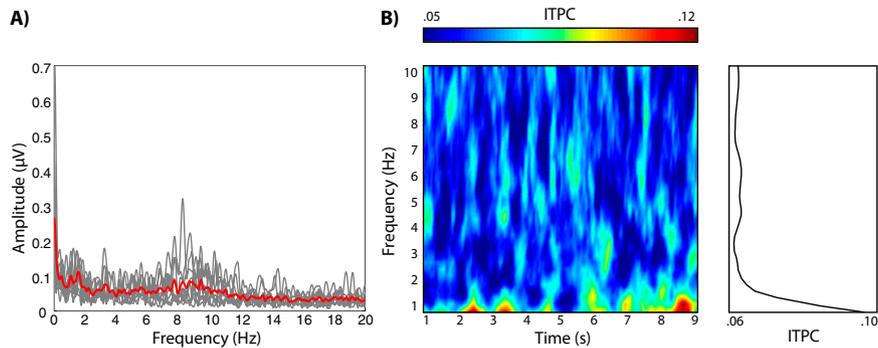
where  $t(j)$  is the time step ( $t = 0-0.33$  s),  $f_m$  was fixed at 3 Hz, and the phase lag parameter,  $\phi$ , was free. Using the best-fit equation, we estimated the time step,  $t(j)$ , at which the function reached a local maximum (hits, P2) or minimum (RT, N1), corresponding to peak performance. The value of  $t(j)$  yielding best performance was then multiplied by  $2\pi f_m$ , where  $f_m = 3$ , yielding the phase angle in radians corresponding to peak performance, i.e., optimal phase. Optimal phases were tested against uniformity by using Rayleigh tests.

Phase lag parameters for behavioral and electrophysiological data with respect to the stimulus and with respect to the brain were maintained and treated as estimates of the stimulus-behavior lag and brain-behavior lag, respectively. As outlined in

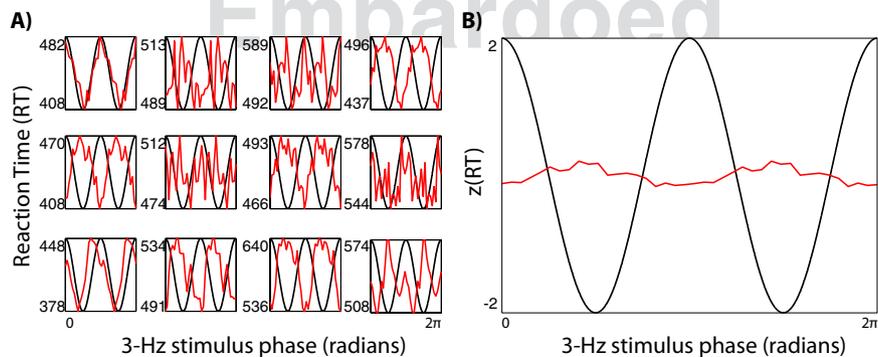
the main text, the combination of the brain–behavior lag and the stimulus–brain lag (estimated from cross-correlations between

the stimulus and time-domain signal) was used to predict the brain–behavior lag.

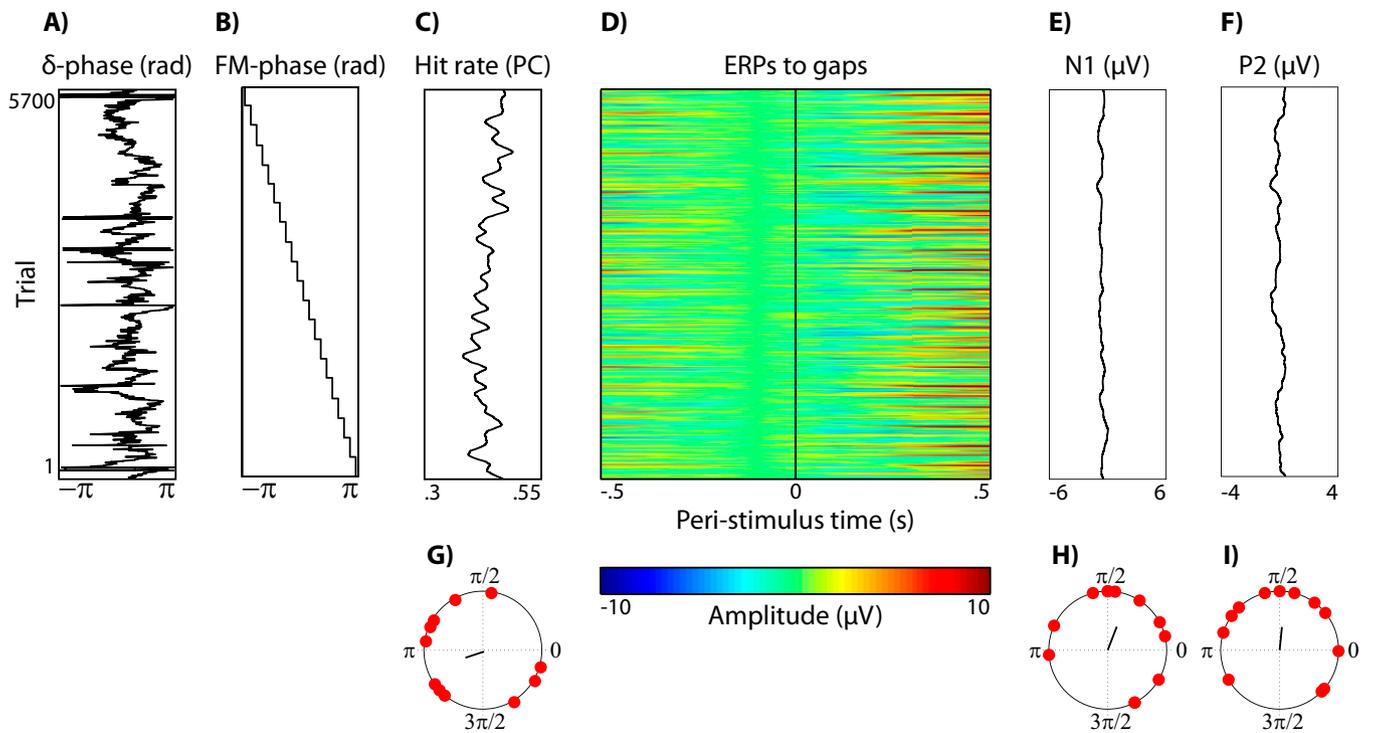
1. Maris E, Oostenveld R (2007) Nonparametric statistical testing of EEG- and MEG-data. *J Neurosci Methods* 164(1):177–190.
2. Busch NA, Dubois J, VanRullen R (2009) The phase of ongoing EEG oscillations predicts visual perception. *J Neurosci* 29(24):7869–7876.
3. Leek MR (2001) Adaptive procedures in psychophysical research. *Percept Psychophys* 63(8):1279–1292.
4. Levitt H (1971) Transformed up-down methods in psychoacoustics. *J Acoust Soc Am* 49(2):2, 467.
5. Studebaker GA (1985) A “rationalized” arcsine transform. *J Speech Hear Res* 28(3): 455–462.
6. Tallon-Baudry C, Bertrand O (1999) Oscillatory gamma activity in humans and its role in object representation. *Trends Cogn Sci* 3(4):151–162.
7. Tallon-Baudry C, Bertrand O, Delpuech C, Pernier J (1997) Oscillatory gamma-band (30–70 Hz) activity induced by a visual search task in humans. *J Neurosci* 17(2):722–734.
8. Lachaux J-P, Rodriguez E, Martinerie J, Varela FJ (1999) Measuring phase synchrony in brain signals. *Hum Brain Mapp* 8(4):194–208.



**Fig. S1.** No evidence for entrainment was observed when amplitude spectra and ITPC were calculated for trials that were not realigned to a common stimulus phase. (A) Amplitude spectrum of FFT of time-domain EEG signal. Red line indicates the group average spectrum, and gray lines show single participants' spectra, averaged over all electrodes. Individual trials were time locked to the stimulus onset and were not realigned with respect to per-trial stimulus phase. Amplitude in the 3-Hz and 6-Hz frequency bins did not differ significantly from amplitude in the neighboring bins [3 Hz:  $t_{(11)} = -1.91$ ,  $P = 0.08$ ; 6 Hz:  $t_{(11)} = -1.71$ ,  $P = 0.12$ ], and the trend was in the wrong direction. (B) ITPC shown over time (Left) and averaged over time (Right), again averaged over all electrodes. A permutation  $t$  test on ITPC (Fig. S2B) failed to reveal any frequency bands in which ITPC was significantly higher than during the prestimulus baseline period.



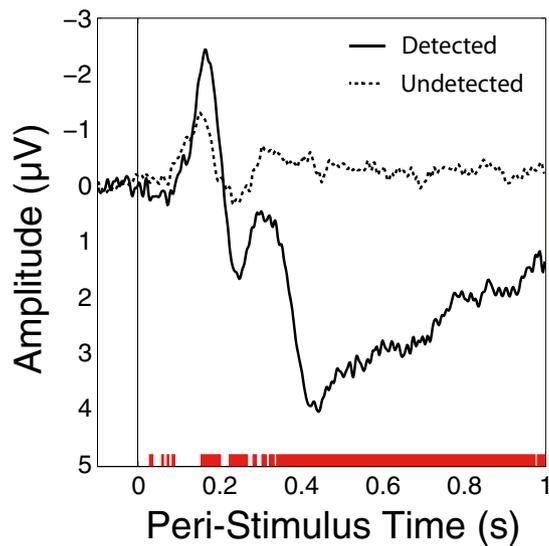
**Fig. S2.** (A) RTs to detected gaps were modulated by stimulus phase. Squared, arcsine-transformed, circular-linear correlation coefficients were calculated for each individual and tested against the null hypothesis of zero correlation. RTs were significantly correlated with stimulus phase [rms  $\rho = 0.70$ ,  $t_{(11)} = 6.74$ ,  $P < 0.001$ ]. (B) However, grand average RTs (z-transformed before averaging) were not systematically related to stimulus phase, ruling out an acoustic explanation for the behavioral modulation.



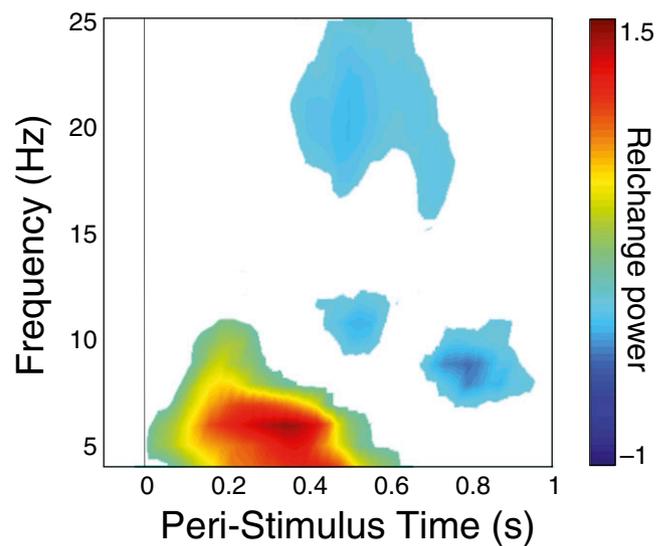
**Fig. S3.** No systematic relation between stimulus phase and hit rate or ERPs was observed across listeners. Trials were sorted according to single-trial stimulus phase (B) rather than single-trial delta brain phase (A). Neither hit rate (C) nor ERPs (D–F) were systematically related to stimulus phase across listeners. Hit rates (C), N1 amplitude (E), and P2 amplitude (F) were each significantly correlated with stimulus phase within listeners. However, optimal stimulus phase defined in terms of hit rate (G), N1 amplitude (H), and P2 amplitude (I) was not consistent across listeners (Rayleigh's tests, all  $P \geq 0.16$ ).

Embargoed

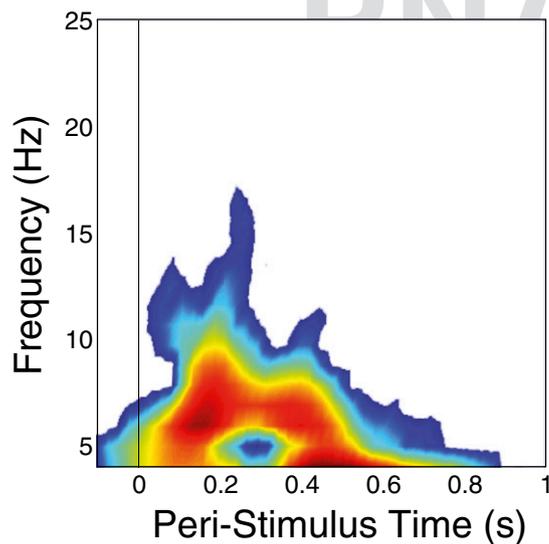
### A) ERPs to gaps



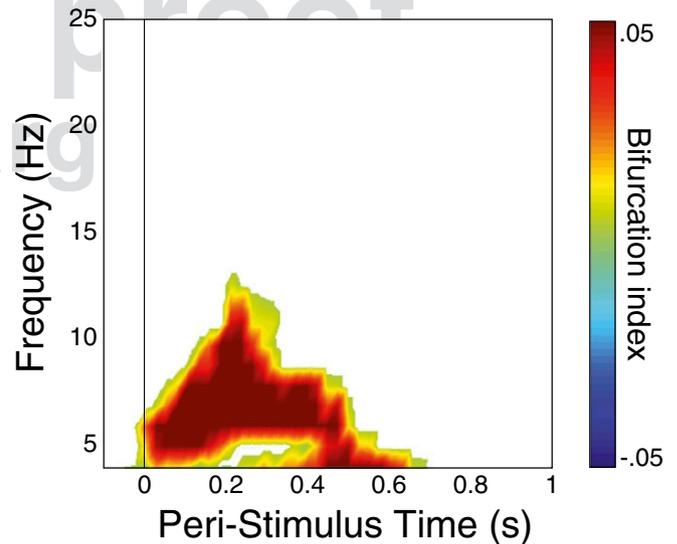
### B) Detected > undetected power



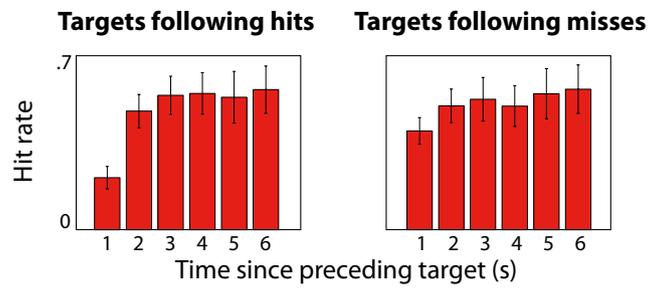
### C) Detected > undetected ITPC



### D) Bifurcation index



**Fig. S4.** Comparison of detected and undetected gaps. (A) ERPs to detected gaps were larger than to undetected gaps; red bar marks areas of significant difference. (B) Significant cluster for total power: detected > undetected gaps. A delta-theta-alpha enhancement during the time window of the ERP was larger for detected gaps, as were later suppressions in the alpha and beta ranges. (C) Significant cluster for ITPC for the contrast detected > undetected gaps. One significant cluster in the delta-theta-alpha range during the time window of the ERP revealed increased phase locking for detected relative to undetected gaps. (D) Significant cluster for the bifurcation index. One significant cluster in the delta-theta-alpha range during the time window of the ERP revealed that detected gaps reset the phase in the 2- to 12-Hz band more consistently than undetected gaps.



**Fig. S5.** Hit rates plotted as a function of the time since the previous gap, separately for cases when the previous gap was a hit (*Left*) versus when it was a miss (*Right*). Hit rates were significantly reduced for targets after a hit within 1 s relative to targets after a miss within 1 s.

PNAS proof  
Embargoed