Title: Test-retest reliability of neural entrainment in the human auditory system

Authors: Yuranny Cabral-Calderin[1*] & Molly J. Henry[1*]

Affiliations: Max Planck Institute for Empirical Aesthetics, Frankfurt am Main, Germany

*** Corresponding authors.**

* Yuranny Cabral-Calderin, Max Planck Institute for Empirical Aesthetics, Grüneburgweg 14, 60322 Frankfurt am Main, Germany. Tel.: (+49) 69 / 8300479 831. Email: yuranny.cabral-calderin@ae.mpg.de

* Molly J. Henry, Max Planck Institute for Empirical Aesthetics, Grüneburgweg 14, 60322 Frankfurt am Main, Germany. Tel.: (+49) 69 / 8300479 830. Email: molly.henry@ae.mpg.de

Abstract: Auditory stimuli are often rhythmic in nature. Brain activity synchronizes with auditory rhythms via neural entrainment, and entrainment seems to be beneficial for auditory perception. However, it is not clear to what extent neural entrainment in the auditory system is reliable over time – a necessary prerequisite for targeted intervention. The current study aimed to establish the reliability of neural entrainment over time and to predict individual differences in auditory perception from associated neural activity. Across two different sessions, human listeners detected silent gaps presented at different phase locations of a 2-Hz frequency modulated (FM) noise while EEG activity was recorded. As expected, neural activity was entrained by the 2-Hz FM noise. Moreover, gap detection was sinusoidally modulated by the phase of the 2-Hz FM into which the gap fell. Critically, both the strength of neural entrainment as well as the modulation of performance by the stimulus rhythm were highly reliable over sessions. Moreover, gap detection was predictable from pre-gap neural 2-Hz phase. Going beyond previous work, we found that stimulus-driven behavioral modulation was better predicted by the interaction between delta and alpha phase than by delta or alpha phase alone, both within and across sessions. Taken together, our results demonstrate that neural entrainment in the auditory system and the resulting behavioral modulation are reliable over time. In addition, both entrained delta and non-entrained alpha oscillatory phase contribute to near-threshold stimulus perception.

Keywords: Reliability, neural entrainment, auditory perception, auditory entrainment, frequency modulation (FM), delta frequency, alpha frequency, oscillations, EEG, phase, brain rhythms.

Introduction

Auditory stimuli, such as music, speech, and animal vocalizations, are often (quasi-)rhythmic in nature. As such, one neural mechanism that has received much recent attention for its contribution to our ability to understand the auditory world is synchronization of brain activity to the rhythms of sounds: neural entrainment (1). Neural entrainment is the process by which neural oscillations phase lock to the rhythms of external sensory stimulation, and has been proposed to be a key mechanism for controlling neural sensory gain (1, 2), attention, and parsing (3) of sensory information that is extended in time. Since neural oscillations are associated with rhythmic fluctuations in the excitation-inhibition balance of neuronal populations (4, 5), neural entrainment by sensory stimulation can modulate perception of physically identical stimuli depending on their timing relative to the phase of entrained neural activity (6, 7). That is, the phase of the entrained oscillation determines whether sensory information is selectively amplified or suppressed (1).

Entrainment to rhythmic environmental stimuli has been described for different sensory modalities (7-9) and across different species (10, 11). In humans, low-frequency M/EEG delta/theta activity has been shown to synchronize to the rhythms of speech and music (quantified as the amplitude envelope of the stimuli), and the success with which this synchronization between brain rhythms and auditory stimuli occurs seems to be critical for successful auditory perception (12, 13), in particular in noisy listening situations (14-17).

If neural entrainment does play a critical mechanistic role for auditory perception, improving the synchrony between brain activity and stimulus rhythms should result in benefits for perception (during e.g., listening in noise). In fact, recent work has used transcranial electrical stimulation (TES) with alternating current (tACS) to interfere with entrainment to auditory stimuli and reported significant modulation of speech comprehension and stream segregation (18-22). For example, applying tACS with speech envelope has been shown to modulate the intelligibility of speech in noise, depending on the phase lag between the electrical stimulation and the speech (21). However, there is some uncertainty in tACS results (see e.g., (23)) and its effectivity for modulating neural entrainment is still under debate. One untested prerequisite for effectively using TES as a targeted intervention is the reliability of neural entrainment over time, i.e., across sessions and days. Establishing the reliability of neural entrainment in the auditory system is a necessary step towards understanding the role of neural oscillations and entrainment for perception, and is critical to pave the way for therapeutic applications based on TES or other noninvasive techniques that target the relationship between neural and

environmental rhythms. The current study aimed to quantify the reliability of neural entrainment in the auditory system over time, and moreover to predict individual differences in auditory perception from neural activity.

Not all stimuli are rhythmic. Thus, multiple neural processing modes have been proposed: "rhythmic-mode" and "continuous-mode" processing (24, 25), only the former of which relies on neural entrainment. Even in the case of purely rhythmic stimulation, lapses of attention have been related to lapses of entrainment, and are rather associated with periods of high-amplitude alpha oscillations (2). Higher alpha amplitude has also been related to reduced entrainment-driven behavioral modulation (26). Taken together, these findings suggest that alpha activity and entrainment represent opposing neural strategies (i.e., continuous-mode vs. rhythmic-mode, or internally vs. externally oriented, respectively), which comodulate behavior. Here, in addition to examining the influence of entrained neural activity on auditory perception (6, 27), we attempted to take the influence of alpha activity into account, in order to provide a more complete picture of the neural mechanisms underlying auditory perception in a vigilance task utilizing rhythmic stimuli.

We employed a paradigm previously used by (7), where stimulus periodicity was communicated by frequency modulation (FM). Participants detected brief auditory targets (silent gaps) embedded in an ongoing 2-Hz FM stimulus (**Fig. 1**). Each participant took part in two EEG sessions. Based on previous literature, we expected that the 2-Hz FM stimulus would entrain delta oscillations in the brain and, as such, gap detection would be modulated by the phase of both the stimulus and the entrained neural oscillation in which target gaps occurred. Furthermore, if neural entrainment is reliable over time, both FM-stimulus induced behavioral modulations and EEG activity should show high inter-session correlations. Our multi-session, within-subject design provided us with a novel opportunity for testing the reliability of entrainment, while controlling for stable individual differences such us anatomical variability or hearing ability. We argue that, while neural entrainment and stimulus-driven behavioral modulation are indeed highly reliable between sessions, auditory perception cannot be exclusively explained by entrained low-frequency neural activity. Both entrained delta phase and non-entrained alpha activity contributed to stimulus-induced modulation of auditory perception.

Results

In two different sessions, EEG activity was recorded while listeners detected silent gaps embedded in 20-s long complex tones that were frequency modulated at 2 Hz (**Fig. 1**). Based on previous literature, we predicted that delta oscillations would be entrained by the 2-Hz FM. We aimed to quantify how reliable that entrainment would be across sessions, both in terms of entrainment strength and the phase relationship between stimulus and brain.

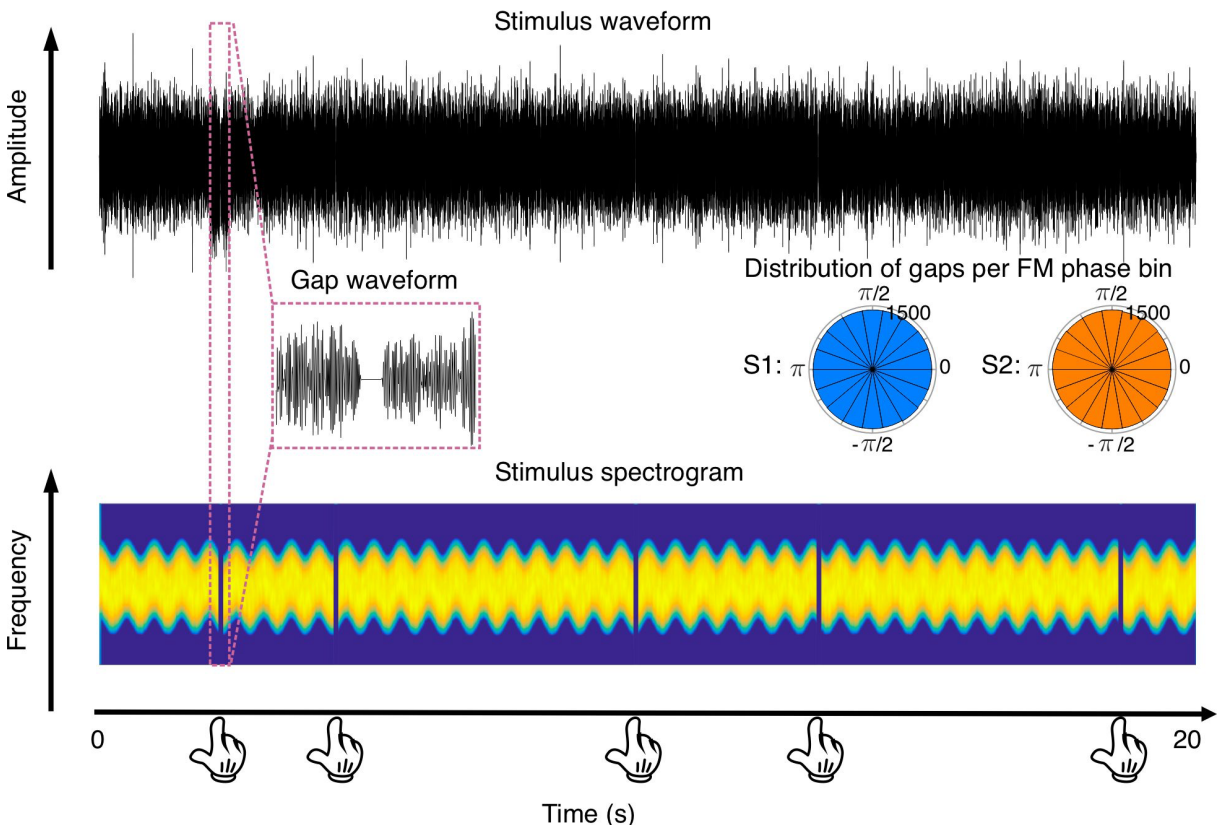

Fig. 1. Auditory stimuli. Stimuli were 20-s long frequency modulated (FM) sounds whose frequency fluctuated rhythmically at 2 Hz (bottom), without periodic fluctuations in amplitude (top). Participants detected short silent gaps (middle left: Gap waveform) embedded in the sound in one of 18 possible phase bins, uniformly distributed around the 2 Hz FM cycle. Each sound had 3, 4, or 5 gaps. Participants responded with a button press each time they detected a silent gap. Circular histograms in the figure (middle right: Distribution of gaps per FM phase bin) show the distribution of gaps per phase bin across participants, separated by session. S1: session 1; S2: session 2. Hand icon was downloaded from https://www.stockio.com/.

Auditory entrainment to FM sounds has high inter-session reliability

We evaluated entrainment using four converging analyses. We considered both total and evoked amplitudes of EEG data. In addition, we calculated inter-trial phase coherence (ITPC)

across the full epoch based on complex Fourier output, and in a time-resolved way based on the output of a wavelet convolution. For evoked amplitude spectra, single-trial time-domain data were first averaged over trials, and then subjected to a Fast Fourier transform (FFT); this analysis is particularly sensitive to neural activity that was phase-locked to the stimulus rhythm across trials. For the total amplitude spectra, the FFT was computed for each single trial, and then the single-trial frequency-domain signals were averaged; this analysis is mostly sensitive to high-amplitude neural activity, and is agnostic to whether that activity is phase-locked across trials. Finally, ITPC was calculated based on the phase (rather than the amplitude) information from the complex output of a FFT calculated on the full-stimulus epoch (with onset- and offset-responses removed) or using the time-resolved complex output of the wavelet convolution (see Methods); these analyses quantify trial-to-trial consistency (i.e., resultant vector length) but are mostly insensitive to amplitude information (28).

For evoked spectra, we observed high-amplitude peaks at the stimulus FM frequency and its first harmonic (i.e., 2 Hz and 4 Hz respectively, **Fig. 2a**), consistent with neural tracking of the rhythm of the FM stimulus (6, 7). Relatively high spectral amplitude was also observed in the alpha frequency band (7-12 Hz). For all further analyses on alpha activity, we considered activity between 7 Hz and 12 Hz, because this frequency range best captured the observed increase in alpha evoked amplitude (**Fig. 2a**). To test for statistical significance, evoked amplitudes for 2 Hz, 4 Hz, and alpha (averaged over 7–12 Hz) frequencies, averaged over all electrodes, were compared to the average amplitude of the neighboring frequency bins (i.e., ±8 frequency bins/0.16 Hz for 2 Hz and 4 Hz and ±100 frequency bins/2 Hz for alpha, see Methods) similar to previous studies (7, 27). Evoked amplitudes for 2 Hz, 4 Hz, and alpha frequencies were significantly different than the average amplitude of the neighboring frequency bins (all $p < 2.95e-07$, see Methods). Total amplitude spectra showed high amplitude in alpha frequency band, while 2-Hz and 4-Hz amplitudes were less visible in the spectra due to the high 1/f power (**Fig. S1**). Nevertheless, compared to the average amplitude of the neighboring frequency bins, 2 Hz, 4 Hz, and alpha amplitudes were also significant in the total amplitude spectra (all $p < 0.001$). Moreover, the ITPC analysis showed clear peaks at 2 Hz and 4 Hz, again suggesting entrainment at the stimulus FM frequency and its first harmonic (**Fig. 2b**). In both sessions, ITPC at 2 Hz and 4 Hz was significantly different than the neighboring frequency bins (same neighboring frequency bins as defined for the evoked amplitude analysis, all $p < 7.74e-08$). Finally, time-resolved ITPC at 2 Hz and 4 Hz, averaged over electrodes and time was significantly different than for the neighboring frequency bins (±8 frequency bins/0.81Hz, all $p < 1.07e-07$, **Fig. 2b**, right). As observed in **Fig. 2a**, FM-stimulus-evoked

amplitude at 2 Hz and 4 Hz was mostly observed in a fronto-central cluster including electrodes F3, Fz, F4, FC1, FC2, C3, Cz, C4, F1, F2, FC3, FC4, C1 and C2 (**Fig. 2a**, insets). Therefore, all further analyses involving these frequencies were done first independently by electrode and then averaged over this subset of electrodes.

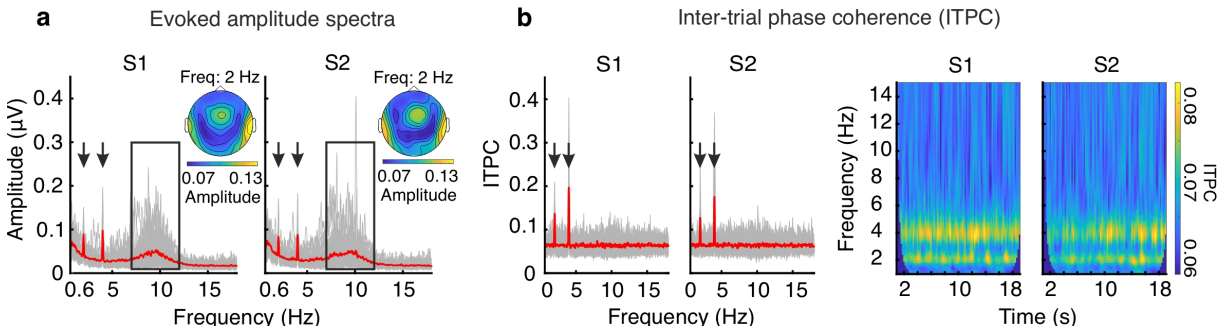

Fig. 2. Neural entrainment to 2-Hz FM stimulus. (a) Evoked amplitude spectra from the fast Fourier transform (FFT) of the time-domain EEG signal. Red solid lines indicate the group average spectrum, gray lines show single participants' spectra, averaged over all electrodes. Inset plots show the topography for the 2 Hz amplitude spectrum averaged across participants, separately for session 1 and session 2. (b) Inter-trial phase coherence (ITPC) indicating the degree of phase clustering across trials for each frequency (left), averaged over all electrodes. Gray lines show individual values and red lines show group average. (right) ITPC shown over time, again averaged across participants and all electrodes. S1: session 1; S2: session 2. Arrows in (a) and (b) indicate the peaks in amplitude and ITPC at the 2-Hz FM stimulus frequency and its first harmonic. Rectangles in (a) indicate the frequency range considered for further alpha analyses.

Moving a step past previous literature, we next asked whether FM-induced entrainment is reliable over time by correlating the amplitude of the stimulus-evoked activity at 2 Hz, 4 Hz, and in the alpha frequency band, as well as the stimulus-brain lag (i.e., the phase angle of the 2-Hz complex output of the FFT calculated for the full stimulus epoch) across sessions. Inter-session correlations were high and significant for evoked amplitudes at 2 Hz (rho = 0.64, p = 2.71e-05), 4 Hz (r = 0.77, p = 1.21e-08), and in the alpha band (r = 0.78, p = 5.47e-09; **Fig. 3a**). No significant difference was observed in the amplitude spectra between sessions for any of the frequencies of interest. Despite individual variability, stimulus–brain phase lags were not uniformly distributed as reported in previous work, but were significantly phase clustered for both sessions (Rayleigh test, session 1: z = 23, p = 9.52e-13; session 2: z = 17, p = 8.78e-09, **Fig. 3b**). Moreover, phase lags were reliable across sessions as indexed by the high circular–circular correlation (rho = 0.62, p = 0.004) and a circular distance between sessions clustered around zero (Rayleigh test, z = 25.72, p = 1.19e-14). Taken together, amplitude spectra and

phase lags suggested that neural entrainment to FM stimuli is reliable across sessions, which is the first vital prerequisite for targeted interventions of auditory-cortex neural oscillations.

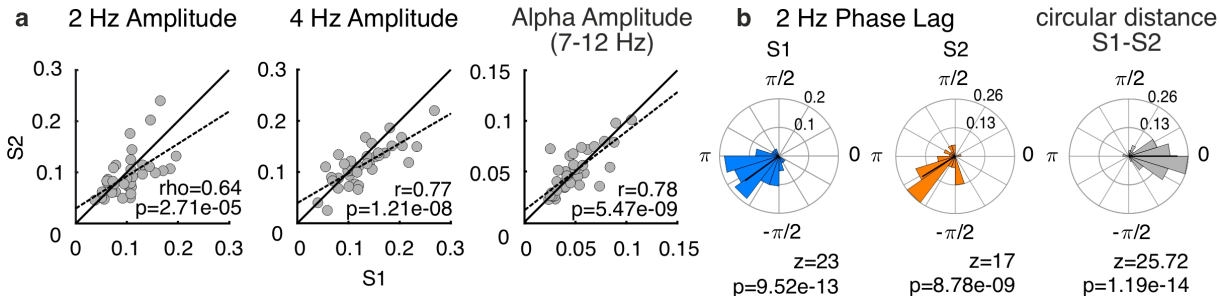

Fig. 3. Reliability of neural entrainment. (**a**) Inter-session correlation of 2 Hz (left), 4 Hz (middle) and alpha (right) amplitudes. Correlation coefficients and associated p-values are given in each plot. Each dot represents a single participant. The solid black line is the diagonal and the dashed line represents the best-fit straight line. (**b**) Circular histograms show neural phase lag relative to the 2-Hz FM stimulus for session 1 (blue, left) and session 2 (orange, middle). The z-values and associated p-values from the Rayleigh test are given in each plot. Circular distance between phase lags in the different sessions is shown in the circular histogram in the Right. S1: session 1; S2: session 2.

Stimulus-induced behavioral modulation is sinusoidal and shows high inter-session reliability

While listening to the FM sounds, participants responded with a button press each time they detected a silent gap. Each 20-s long stimulus contained three, four, or five gaps. Gaps were distributed uniformly around the 2-Hz FM cycle in 18 possible phase positions (**Fig. 1**). A response was considered to be a "hit" if a button press occurred within a window of 0.1-1.5 s after gap onset.

We hypothesized that, as a consequence of the stimulus-induced entrainment of delta oscillations, hit rates for gap detection would be sinusoidally modulated by the FM stimulus phase. To go a step past previous studies, we also asked whether this modulation is reliable across sessions. While several studies have focused on analyzing oscillatory modulation of perception at the group level by aligning single-participant data to the phase with best or worst performance (see (29) for recommendations), here we took a different approach and investigated whether the magnitude of stimulus-driven behavioral modulation and the preferred (best) FM-stimulus phase were consistent across sessions within an individual.

For each session separately, hit rates were calculated for each FM-phase bin (**Fig. 4a, Fig. S2**). Then, we fit a cosine function to hit rates as a function of phase for each participant and each

session. The resulting amplitude parameter from the cosine fit was taken as an index of the strength of the behavioral modulation. Significance of the sinusoidal modulation was tested using a permutation approach, whereby 1000 surrogate datasets were created for each participant and session by shuffling the single-gap accuracy values (0,1) with respect to their stimulus-phase labels. Cosine functions were fitted to the surrogate data and the fitted amplitude from the real data was compared for each participant and session to the surrogate data (see Methods). Significant behavioral modulation was observed for 32/38 participants in session 1 and 37/38 in session 2 (z-score ≥ 3.2, $p < 0.05$, Bonferroni corrected, **Fig. 4b**). Similar sinusoidal modulation was also observed for the reaction times but with opposite phase lag (i.e., high accuracy and fast RTs occurred in the same 'optimal' stimulus phase; **Fig. S3a**).

The fitted modulation amplitude values were highly correlated between sessions ($r = 0.88$, $p = 2.22e-13$, **Fig. 4b**), indicating that FM-induced behavioral modulation was reliable. Individual preferred FM phases (the FM-stimulus phase yielding highest performance) were estimated from the fitted cosine functions per participant and per session (**Fig. 4b**). Preferred phases were not uniformly distributed, but clustered in one half of the FM cycle (Rayleigh test; session 1: $z = 34.4$, $p = 1e-22$; session 2: $z = 35.7$, $p = 1.76e-24$). Preferred phases were also highly correlated between sessions (circular–circular correlation; rho = 0.62, $p = 0.036$, **Fig. 4b**). Moreover, circular distance between preferred phases in session 1 and session 2 was clustered around zero (Rayleigh test; $z = 35.7$, $p = 1.94e-24$), suggesting that preferred phase is also a reliable attribute of FM-induced behavioral modulation.

We recognized that high reliability in preferred FM phase across sessions may have been at least partially attributable to clustering of preferred phases across participants in the first place – since preferred phases were significantly clustered across participants, we wanted to be careful not to overinterpret similar mean preferred phase across sessions as reflecting within-participant reliability. In order to test this possibility, we conducted a permutation test in which the order of participants in session 2 was permuted 1000 times relative to the session-1 order, and the circular–circular correlation between individual preferred phases in session 1 and (permuted) 2 was computed. If inter-session correlations are driven by the similarity between participants, permuting the participant order in session 2 should not have affected inter-session correlations. However, demonstrating the reliability of preferred phase, the inter-session correlation in the original data was significantly higher than inter-session correlations on the permuted samples ($z = 3.66$; $p < 0.001$, **Fig. S3b**).

In addition to the fitted amplitude parameters and preferred FM phases for gap-detection performance, hit rates (over the entire session), false-alarm rates, reaction times, and threshold gap durations were also highly correlated across sessions (Rho ≥ 0.5, $p \leq 0.002$, **Fig. 4c**). When each of the dependent measures were directly compared across sessions, a significant difference was observed only for the gap duration (*Wilcoxon signed rank test*, z = 2.99, p = 0.003). Individually adjusted gap durations estimated using our threshold procedure were shorter in session 2 than in session 1. Since hit rates were not significantly different between sessions, the decrease in threshold gap duration suggests that participants experienced some degree of learning or practice effect, and could recognize shorter gaps in session 2. The reduction in threshold gap duration between session 1 and 2 was significantly correlated with music perceptual abilities (Rho = 0.39, p = 0.017), as measured with the Goldsmiths Musical Sophistication Index (Gold-MSI, (30)), very tentatively suggesting that individuals with stronger music skills might have experienced a greater benefit of repeated exposure across session.

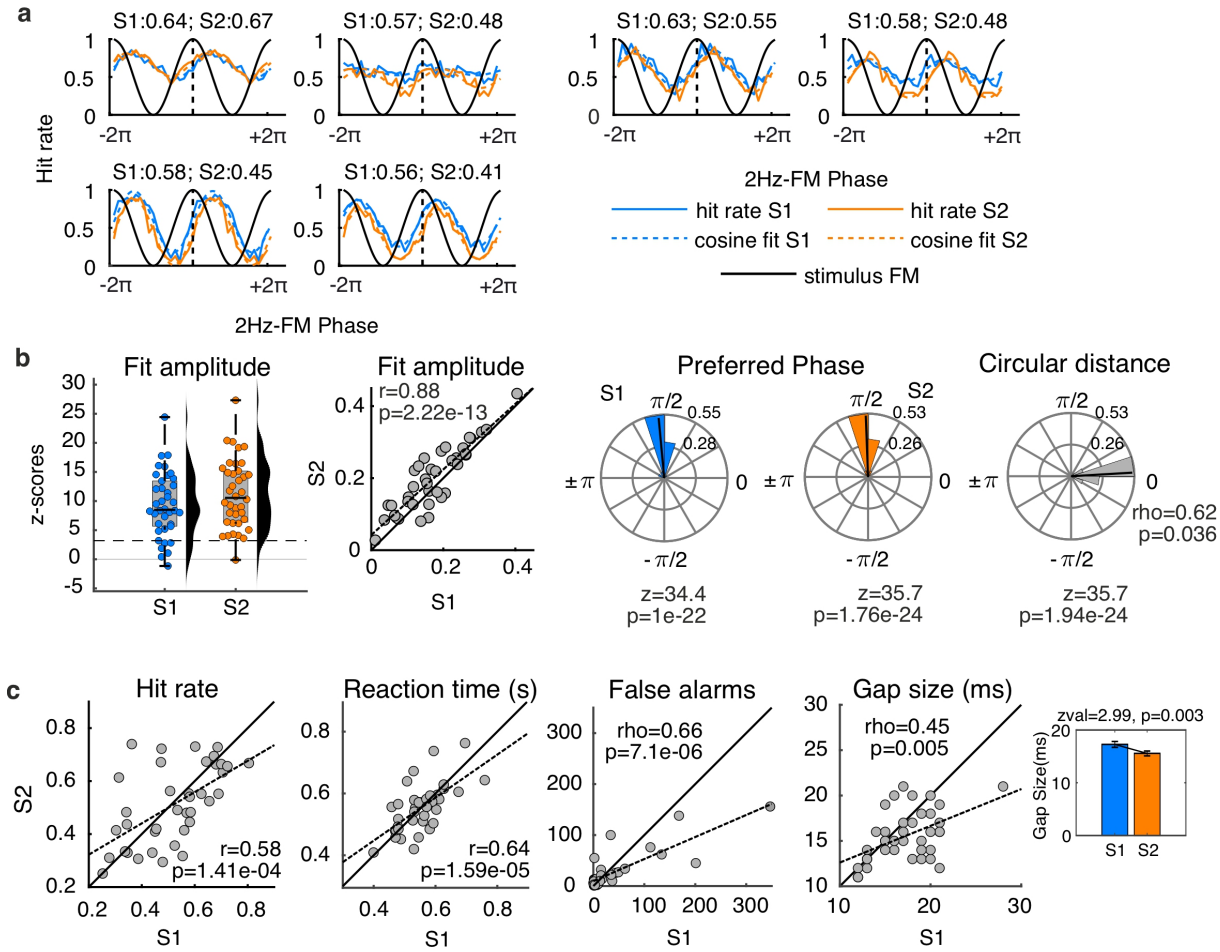

Fig. 4. Stimulus-driven behavioral modulation and its reliability. (**a**) Hit rates as a function of 2-Hz FM stimulus phase. Dashed blue and orange lines represent the fitted cosine functions for sessions 1 and 2, respectively. Numbers on top of each graph show the mean hit rate across phase bins for each session. Each graph shows data for a different single participant. In all panels, blue represents session 1 (S1) and orange represents session 2 (S2). (**b**) The plot on the left shows the distribution and spread of the z-scores of the amplitude of the observed sinusoidal modulation for each session. The box plot shows median (black horizontal line), 25th and 75th percentiles (box edges) and extreme datapoints (whiskers). Each circle represents a single participant. Horizontal dashed line marks the significance threshold of z = 3.2 (*Bonferroni* corrected z-score value for 76 comparisons -i.e., 38 subjects x 2 sessions). Scatter plot in the middle shows the correlation between the fit amplitudes for the 2 sessions. Solid line is the diagonal and dashed line shows the best-fit straight line. Circular histograms on the right of the panel show individual preferred phases (i.e., phase in the cosine fit with highest hit rate) separated by session (left and middle histograms) and the circular distance between the two (right histogram). *Z* and *P-values* in the plots refer to the results from the Rayleigh test. (**c**) Scatter plots show inter-session correlations for hit rates, reaction times, false alarms rates, and gap durations. Solid lines are the diagonal and dashed lines show the best-fit straight lines.

Pre-stimulus neural 2-Hz phase predicts gap detection

In the previous sections, we showed that the 2-Hz FM stimulus entrained neural activity at the modulation frequency and that gap-detection performance was sinusoidally modulated by FM phase. Therefore, we expected that pre-gap brain activity should also predict gap-detection performance. We examined the effects of the neural phase and amplitude in the FM stimulus frequency band (2 Hz), as well as the neural phase and amplitude in the alpha frequency band (7-12 Hz), extracted from the pre-gap time window (see Methods). Both entrained 2-Hz and non-entrained alpha activity were taken into account since our initial FFT analysis showed stimulus-driven modulation of both (**Fig. 2a**). Five logistic regression models were fitted to the individual data using different combinations of regressors aiming to predict trial-based gap detection performance (hit/miss). For model selection, Akaike's information criterion (AICc, corrected for small samples) values were averaged across participants separated by session (**Table S1**). In both sessions, the smallest AICc values were obtained for the model that included pre-gap neural 2 Hz phase, pre-gap neural alpha phase and pre-gap alpha amplitude. Note however, that the difference in AICc values between models was small in some cases. To further test the significance of each predictor at the group level, the individual beta estimates were compared to the mean beta estimates obtained from fitting the same model to surrogate datasets (see Methods). For each subject and session, 1000 surrogate datasets were created by shuffling the dependent variable, single-gap accuracy values (0,1), while keeping all independent regressors fixed.

Gap detection was significantly modulated by pre-gap 2-Hz phase (S1: $t(37) = 3.31$, $p = 0.01$; S2: $z = 3.14$, $p = 0.01$, **Fig. 5a**). No significant effect of alpha phase was observed (uncorrected $p > 0.16$, S1 and $p > 0.38$ S2, **Fig. 5a**). The effect of alpha amplitude was significant although

it did not survive *Bonferroni* correction (S1: z (37) = 2.17, p = 0.03, uncorrected; S2: t (37) = 2.2, p = 0.03, uncorrected, **Fig. 5a**). Note that although alpha phase and amplitude were not considered significant predictors of gap-detection performance using this permutation strategy, they did contribute to the best-fitting logistic regression model, as indexed by AICc. No significant difference was observed between the two sessions for any of the predictors.

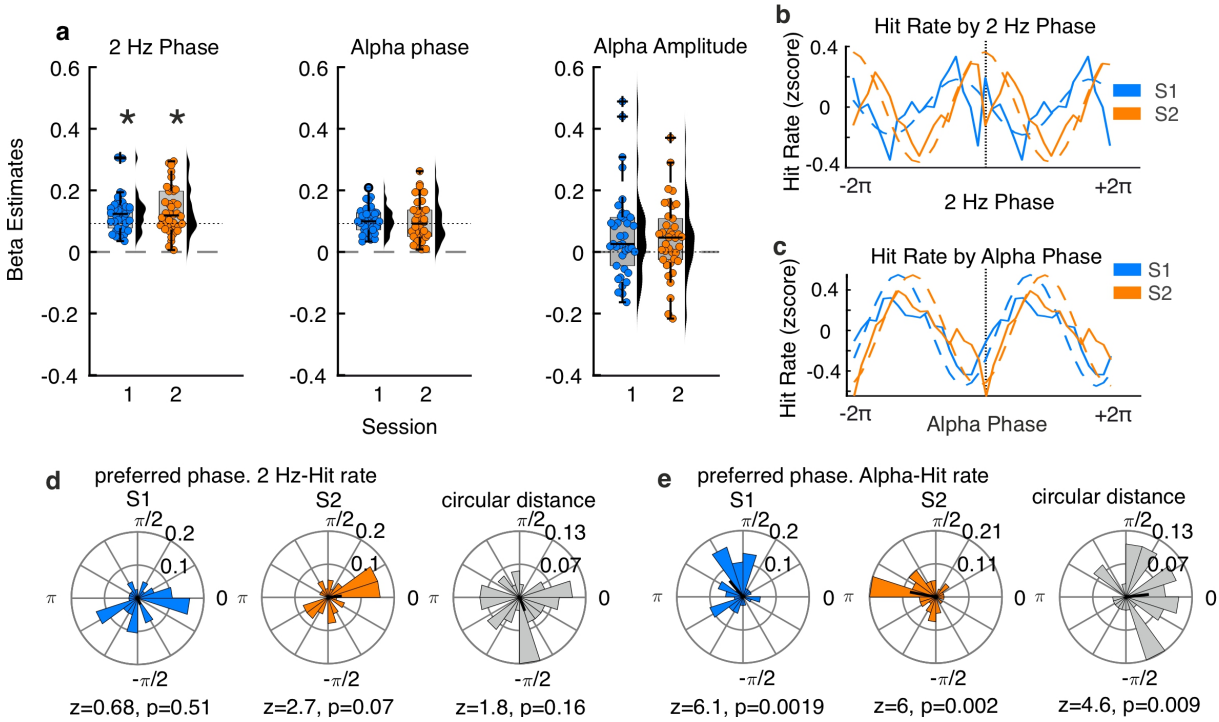

Fig. 5. Effect of pre-gap activity on gap detection. (**a**) Beta estimates (including distribution and spread) for 2-Hz phase (left), alpha phase (middle), and alpha amplitude (right) from the winning individual logistic regression models fitted to the EEG data. Box plots show median (black horizontal line), 25th and 75th percentiles (box edges) and extreme datapoints (whiskers). Black crosses represent outlier values. Each circle represents a single participant. Horizontal dotted lines mark the median of the beta estimates obtained from the logistic regressions fitted to the surrogate datasets, against which the beta estimates from the true data were compared. P-values are *Bonferroni* corrected and show the significant effect at the group level for a given predictor. (**b**) Effect of 2-Hz pre-gap phase. For visualization, the pre-gap neural phase at 2 Hz was grouped in 18 equally spaced bins and hit rates were calculated for each bin according to its pre-gap 2-Hz phase. The figure shows the average across participants. Two cycles are shown for illustration purposes. Solid lines represent the actual data and dashed lines represent the fitted cosine function. (**c**) Similar to (**b**), but the data were binned according to pre-gap alpha phase. (**d**) Preferred neural 2-Hz phase separated by session (left, middle) and circular distance between the two sessions' preferred phases (right). (**e**) Same as in (d) but for pre-gap alpha phase. S1: session 1 (blue); S2: session 2 (orange).

Next, we evaluated the clustering across participants of preferred neural 2-Hz and alpha phases and tested the reliability of preferred neural phases across sessions. Trials were sorted and binned according to the instantaneous neural phase at 2 Hz (**Fig. 5b**) or in the alpha frequency band (7-12 Hz, **Fig. 5c**), using 18 equally spaced phase bins. Hit rates were calculated for each

bin and cosine functions were fitted to each individual participant's data in order to estimate preferred neural phase, similar to the behavioral analysis. Rayleigh tests did not show any significant deviation from uniformity for 2-Hz preferred phases within either session (all $Z < 2.7$, $p > 0.0.07$). In line with this lack of consistency, we observed a nonsignificant clustering of circular distances between preferred phases across sessions ($Z = 1.8$, $p = 0.16$, **Fig. 5d**) and a nonsignificant correlation between preferred phases across sessions for the optimal 2-Hz neural phase for gap detection (circular-circular correlation Rho = 0.12, $p = 0.44$). In contrast, Rayleigh tests showed significant clustering of preferred alpha phases across participants in both sessions ($Z > 6$, $p < 0.002$). Moreover, circular distances between preferred alpha phases across sessions were significantly clustered ($Z = 4.6$, $p = 0.009$, **Fig. 5e**), although preferred alpha phases did not significantly correlate across sessions (Rho = 0.24, $p = 0.12$). Thus, while 2-Hz preferred neural phases were randomly distributed across participants and were not stable between sessions, preferred alpha phase was more consistent across participants and sessions. Gap-evoked potentials (ERPs) can be seen in supplemental results and Fig. S5.

Stimulus-driven behavioral modulation is better predicted by the interaction between entrained 2-Hz and non-entrained alpha phase effects than by either frequency band alone

Our analyses up to this point demonstrated that single-trial gap-detection performance was mostly predictable from 2-Hz pre-gap phase. However, in the logistic regression models, 2 Hz and alpha phase effects were modelled independently and therefore no phase–phase interaction between frequency bands was tested. Therefore, we next asked whether single-trial gap-detection performance was co-modulated by pre-gap neural 2 Hz phase and pre-gap alpha phase, which could potentially explain the inconsistency in 2-Hz preferred phases both within and across sessions. Specifically, we next asked whether the stimulus-driven sinusoidal modulation of hit rates could be explained better by entrained 2-Hz phase, by ongoing alpha phase, or the interaction between the two (**Fig. 6**).

For each gap, we determined the pre-gap 2-Hz neural phase and the pre-gap neural alpha phase and assigned each to 1 of 18 bins, as in the analyses described above. Then, 3 detection probabilities were assigned to the gap: 1) the participant-and-session-specific hit rate for the corresponding 2-Hz neural phase bin, 2) the hit rate for the corresponding alpha neural phase bin, and 3) the product of the two hit rates (i.e., the interaction; **Fig. 6a,** see Methods). Gaps were then binned by FM-*stimulus* phase and 3 predicted performance functions were created by averaging the predicted probabilities across all gaps within a FM-stimulus bin for the 3

models based on 2-Hz neural phase ($_{pred}$2Hz), alpha phase ($_{pred}\alpha$), and the interaction ($_{pred}$2Hz*α) predictions (**Fig. 6a**). Then, the true hit-rate profiles as a function of FM-stimulus phase were correlated with the three predicted functions so that we could determine which of the predictions most closely matched the data.

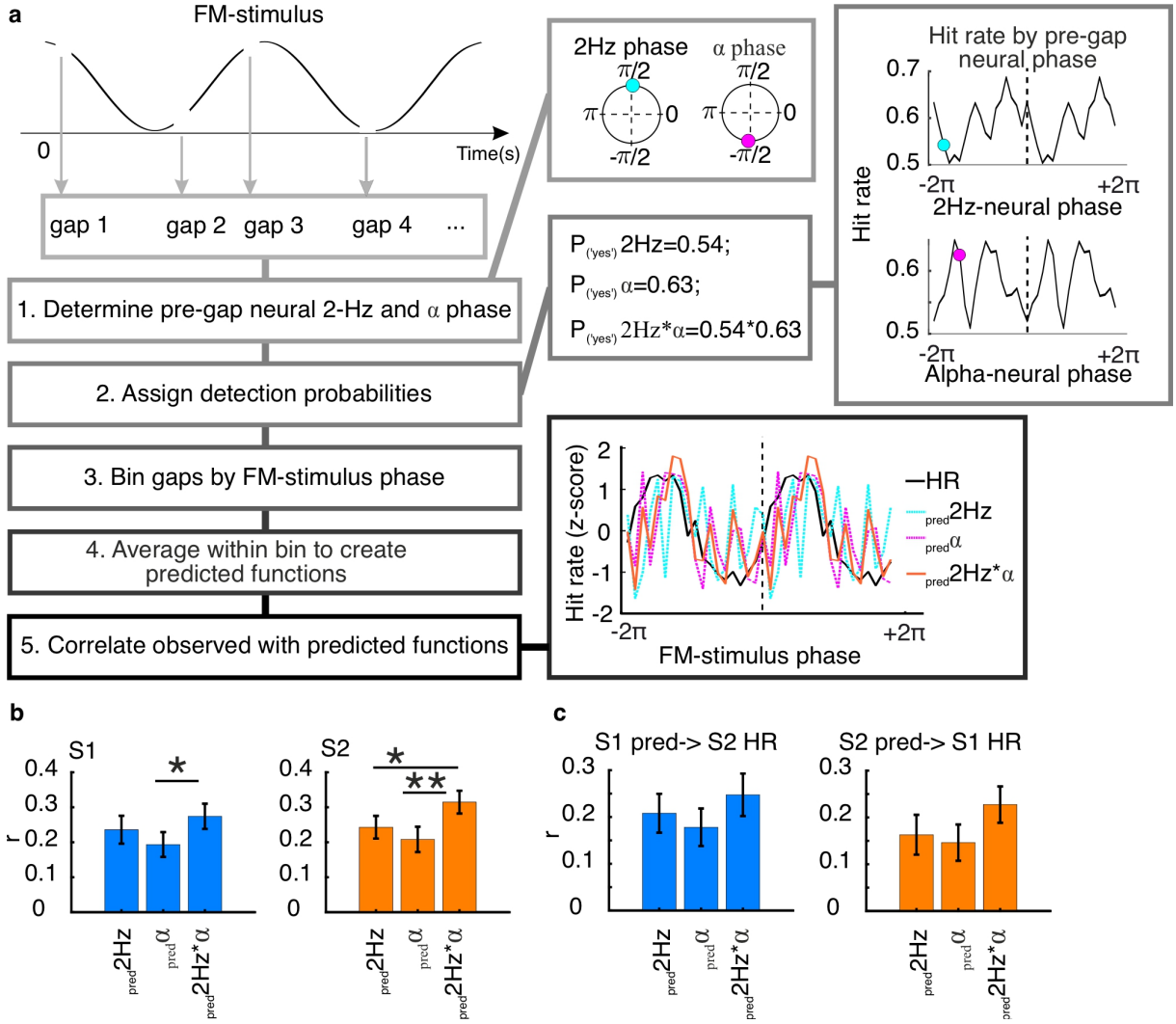

Fig. 6 Predicting FM-stimulus induced behavioral modulation from 2-Hz and alpha neural phase. **(a)** For each gap, pre-gap 2-Hz and alpha neural phases were calculated and detection probabilities ($P_{('yes')}$) were assigned based on 1) the hit rate calculated for the corresponding pre-gap neural 2-Hz phase bin, 2) the hit rate calculated for the corresponding pre-gap neural alpha phase bin, and 3) their interaction (multiplication). Circle colors (magenta and cyan) denoting the pre-gap neural phase for the specific gaps in step 1 correspond to the same bins marked in the average hit-rate plots in step 2. For each FM-stimulus phase bin (18 bins), three predicted functions were calculated by averaging detection probabilities across gaps using the 1) 2-Hz neural phase predictions, 2) the alpha neural phase predictions and 3) their interaction. (**b**) For each participant and session, the true hit-rate profiles as a function of FM-stimulus phase (HR) were correlated with the three predicted functions ($_{pred}$2Hz, $_{pred}\alpha$, $_{pred}$2Hz*α) so that we could determine which of the predictions most closely matched the data. Bar graphs show the mean correlation coefficients across participants. Error bars denote standard error of the mean. * p<0.05, ** p<0.01, *Bonferroni* corrected. Note that the plots show Pearson's correlation coefficients but statistical comparisons where performed on the fisher's r-to-z transformed values. (**c**)

same as in **(b)** but correlation coefficients were computed by correlating the individual hit rate by FM-stimulus phase bin observed in session 2 with the functions predicted in session 1 (S1 pred→ S2 HR, left) and by correlating the individual hit rates as a function of FM-stimulus phase bin observed in session 1 with the functions predicted in session 2 (S2 pred→ S1 HR, left).

In both sessions, observed FM-stimulus-driven modulation of hit rates significantly correlated with all predicted functions (one sample t-test on Fisher's r-to-z transformed correlation coefficients from true vs. mean correlation coefficients from surrogate data, all $t(37) > 5.34$ $p < 2.93e\text{-}05$, *Bonferroni* corrected for 6 comparisons, **Fig. 6b**). No significant differences were observed between sessions for any of the predictors (uncorrected $p > 0.37$). In general, correlation values were higher for $_{pred}2Hz*\alpha$ than for $_{pred}2Hz$ or $_{pred}\alpha$ alone, and this difference was significant for session 1 for the comparison $_{pred}\alpha$ vs. $_{pred}2Hz*\alpha$ ($t(37) = -3.12$, $p = 0.02$, *Bonferroni* corrected for 6 comparisons) and session 2 for $_{pred}2Hz$ vs $_{pred}2Hz*\alpha$ ($t(37) = -2.83$, $p = 0.04$) and $_{pred}\alpha$ vs $_{pred}2Hz*\alpha$ ($t(37) = -3.82$, $p = 0.002$, all *Bonferroni* corrected for 6 comparisons, **Fig. 6b**).

In addition, we asked whether the modulation of hit rates by 2-Hz and alpha phase from one session could predict the same in the other session. Stimulus-driven behavioral modulation was significantly correlated with the predictor functions from the opposite session for all three predictors (all $t(37) > 3.66$; $p < 0.005$, *Bonferroni* corrected for 6 comparisons, **Fig. 6c**). As before, there was a trend to higher correlation values for the interaction predictor $_{pred}2Hz*\alpha$ compared to $_{pred}2Hz$ and $_{pred}\alpha$ although it did not survive multiple comparisons correction (predictors session 1 to HR session 2: $_{pred}\alpha$ vs. $_{pred}2Hz*\alpha$; $t(37) = -2.01$; $p = 0.02$, uncorrected; predictors session 2 to HR session 1: $_{pred}2Hz$ vs. $_{pred}2Hz*\alpha$; $t(37) = -2.6$; $p = 0.01$, uncorrected $_{pred}\alpha$ vs. $_{pred}2Hz*\alpha$; $t(37) = -2.19$; $p = 0.03$, uncorrected). Taken together, these results show that 2-Hz neural phase and alpha neural phase effects both explain the FM-stimulus induced behavioral modulation. However, the interaction between 2-Hz and alpha phase was the best predictor of behavior. Moreover, behavioral modulation could be predicted regardless of whether predictive functions were derived from the same or the opposite session, which suggests that this is also reliable across sessions.

Control Experiment: Gap detection is modulated by the FM rate.

Here, we found that preferred FM phases were clustered in one half of the FM cycle (although we did observe some variability across participants). This was somewhat surprising because previous studies using the same stimuli and task reported preferred phases to be uniformly

distributed around the FM cycle (7). The one major difference between the two studies was the FM frequency: 3 Hz in the previous study (7) and 2 Hz in the current study. We hypothesized that individual variability in stimulus-driven behavioral modulation might be influenced by the FM frequency: slower FM frequencies (in this case 2 Hz) should be easier to track and therefore participants should synchronize to such stimuli more similarly, while individual differences are more pronounced for higher FM frequencies, which might be more difficult to track and therefore result in phase slips.

To test this hypothesis, in a control experiment, 16 participants performed the same gap-detection task as in the main experiment, but FM frequency varied between blocks and took on values of 1.5 Hz, 2 Hz, 2.5 Hz, or 3 Hz (see Methods, **Fig. S4**) Overall hit rates were significantly modulated by FM rate (repeated measures ANOVA, $F(3,45) = 8.49$, $p = 0.0001$) with smaller hit rates observed for higher FMs (**Fig. S4a**). However, FM rate did not significantly affect the amplitude of the stimulus-induced behavioral modulation ($F(3,45) = 1.72$, $p = 0.18$). Most importantly here, to test whether the clustering of individual preferred phases was dependent on the FM rate, individual preferred phases were estimated from cosine fits **(Fig. S4b**) and the resultant vector length across participants was calculated for each FM rate (**Fig. S4c**). A linear model was fitted to the resultant vector length including intercept and a linear term for FM rate as the predictor. A significant fit for the linear term would suggest that indeed, phase clustering significantly decreased (or increased) with increasing FM rate. Significance was evaluated using permutation tests. The permutation distribution was created by 1) computing resultant vector lengths over 1000 iterations (in each iteration the FM rate labels where permuted for each participant) and fitting the same linear model to the simulated vector lengths as for the original data. The t-value calculated for the effect of FM rate on vector length in the original data was compared to the distribution of t-values obtained with the permuted datasets. Using this approach, no significant effect of FM rate on vector length was observed ($z = -0.52$, $p = 0.3$). After visual inspection of the data, we tested a second linear model incorporating, in addition to the intercept and linear term, a quadratic term (FM rate ^2). Significance of each function was evaluated using permutation tests, as already described for the linear-only model. In brief, t-values obtained for the linear and quadratic terms were each compared to the random t-values distributions obtained, for each term, by fitting the same linear model to the shuffled datasets. The t-values from the linear and quadratic fits to the original data were significant at trend level compared to the simulated permutation distributions (linear fit: $z = -1.61$, $p = 0.054$; quadratic fit: $z = 1.51$, $p = 0.066$).

Gap detection is also influenced by stimulus carrier and time of occurrence

Although we tried to minimize the possibility that the FM phase effects may have unknowingly been driven by acoustic confounds, we nonetheless tested which stimulus characteristics beyond FM phase influenced gap-detection hit rates. Fitting logistic regression models to individual participants data, we examined the extent to which single-trial gap detection accuracy in the main experiment was influenced by 1) FM phase, 2) global experiment time (when the gap occurred over the whole experiment, quantified as the gap number within a session, 1–864), 3) local time (when the gap occurred within a stimulus), 4) the center carrier frequency of the stimulus in which the gap was presented and 5) the interaction of FM phase with global time (quantifying the extent to which stimulus-driven behavioral modulation changed over the course of the experiment).

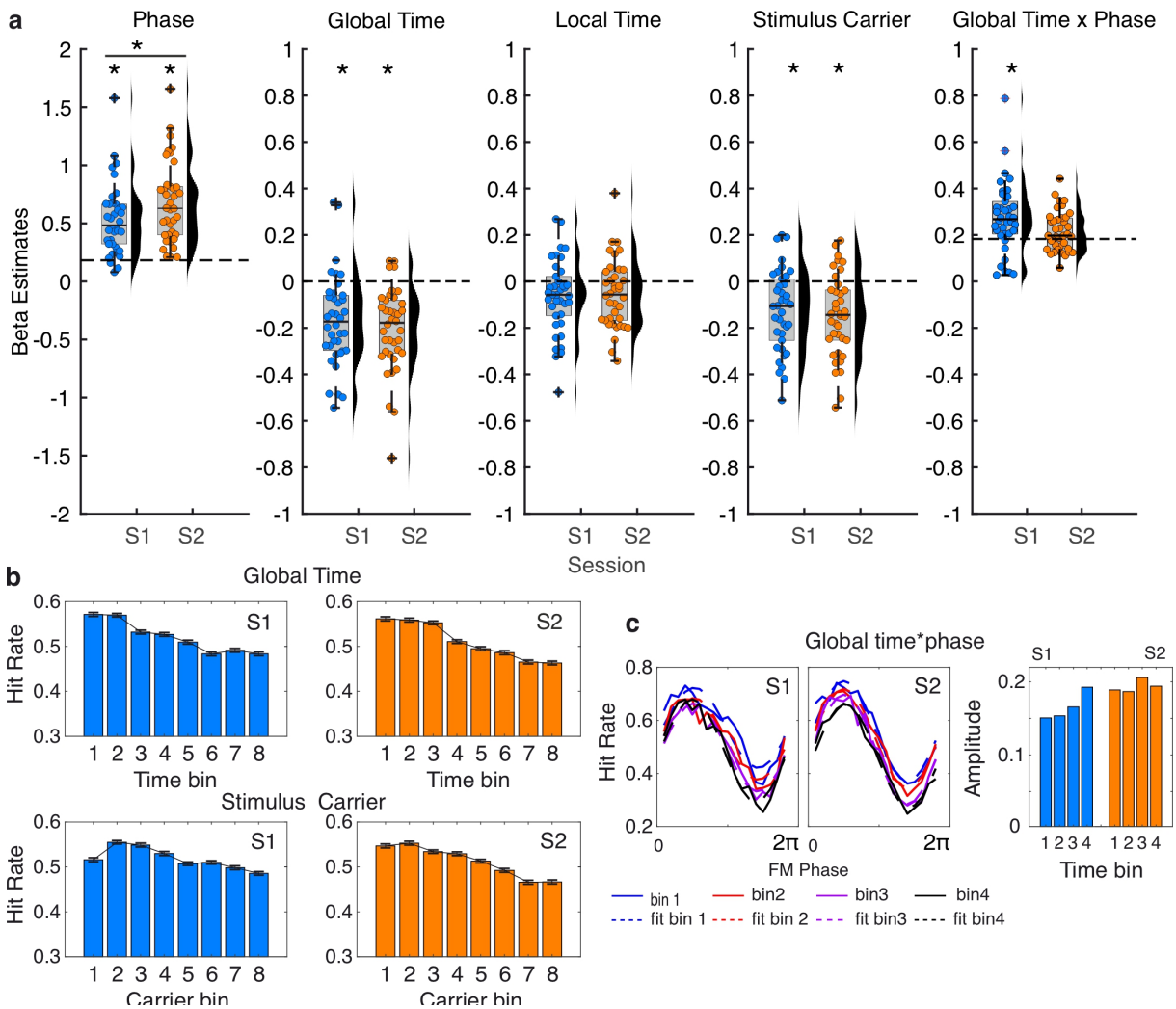

Fig. 7. Effect of stimulus properties on gap detection. (**a**) Distribution plots showing the beta estimates resulting from the individual logistic regression models. (*) denote the significant effect at the group level for a given predictor, after performing one-sample *T-tests* compared to the mean beta estimates for surrogate datasets. (*) denote significant difference between sessions after performing

paired samples T-test. All *<0.05 after *Bonferroni* correction. A significant difference between sessions was only observed for the main effect of phase, which was strongest in session 2. Box plots show median (horizontal solid black line), 25th and 75th percentiles (box border), extreme values (whiskers) and outliers (black cross). (**b**) Visualization of the main effects shown in (**a**), i.e., main effect of global time (top) and main effect of stimulus center carrier frequency (bottom). (**c**) Global time*phase interaction effect. Data were grouped in 4 linearly spaced bins (1: blue; 2: red; 3: purple; 4: black) according to the time of occurrence within the session (global time). Hit rates were calculated for each bin and each FM phase (solid lines, Left). Cosine functions were fitted to each bin data (dashed lines, Left). Mean fit amplitude values are given in the bar plots in the Right. S1: session 1; S2: session 2.

Significance at the group level was estimated using one-sample (significant effect for each session) and dependent-sample (comparison between sessions) t-tests on the individual beta estimates (see Methods). One-sample t-tests were performed comparing the individual beta estimates obtained from the logistic regression on the original data with the mean beta estimates obtained when fitting the same models to surrogate datasets. For each subject, 1000 surrogate datasets were created by shuffling each time the dependent variable, single-gap accuracy values (0,1), while keeping all independent variables (stimulus parameters) fixed.

As already described, in both sessions, gap-detection performance was significantly modulated by FM-stimulus phase (S1: t (37) = 7.04, p = 3.55e-07; S2: t (37) = 7.42, p = 1.1e-07 *Bonferroni* corrected). A significant influence on hit rates was also observed, in both sessions, for global experiment time (S1: t (37) = -5.26, p = 8.89e-05; S2: t (37) = -7.06, p = 3.33e-07, *Bonferroni* corrected) and the center carrier frequency of the stimulus (S1: t (37) = 3.77, p = 0.008, *Bonferroni* corrected). The interaction FM phase*global time was significant only in session 1 (t (37) > 7.84 p < 4.02e-08, *Bonferroni* corrected, **Fig. 7a**) but it did not survive correction for multiple comparisons in session 2. Significant difference between sessions was observed only for the effect of FM phase, which was strongest in session 2 (t (37) = -3.6, p = 0.006, *Bonferroni* corrected). Regarding the main effect of the stimulus center carrier frequency, post-hoc exploration showed that hit rates decreased with increasing carrier frequency (**Fig. 7b**). The main effect of global time was explained by a decreased of hit rates over time, as it would be expected from fatigue (**Fig. 7b**). Further visual exploration of the FM phase*global time interaction term in session 1 showed that while hit rates decreased over time within the session, the amplitude of the sinusoidal modulation increased over time. The later could represent some sort of strategy change happening within the first session, when participants were more naïve to the experiment (**Fig. 7c**). Additional logistic regression models were fitted to the individual data using other combination of parameters (e.g., including the carrier frequency at the exact time the gap was presented, taking into account the sinusoidal modulation) but such models

showed higher corrected Akaike's Information Criterion (AICc) values than the winning model, suggesting that they were less representative of the data (**Table S2**). Interestingly, the model with the highest AICc (and therefore worst performance) was the one including acoustic parameters but excluding FM stimulus phase information, suggesting that in fact, FM phase was the best predictor of gap detection accuracy.

Discussion

In the present study, we investigated the test–retest reliability of neural entrainment and its relevance for auditory perception. Participants detected silent gaps embedded in 2-Hz FM stimuli while EEG activity was recorded in two separate sessions. We showed that: 1) neural activity was entrained by the FM stimuli, and perception was modulated in a sinusoidal manner by both stimulus and brain phase; 2) both entrainment strength and the magnitude of stimulus-driven behavioral modulation were reliable across sessions; 3) pre-gap delta and alpha activity predicted moment-by-moment fluctuations in performance, within and across sessions.

Neural entrainment is reliable over time

Over the past years, a great deal of attention has been paid to entrainment, and arguments range from the existence of entrainment per se (1, 31, 32) to its role for auditory perception specifically (33-35). Neural entrainment to rhythmic stimuli has been observed across a range of sounds, including simple stimuli such as tone sequences, amplitude-modulated, or frequency-modulated sounds, and more complex auditory stimuli such as music and speech (6, 7, 9, 10, 12, 13, 19, 26). However, until now, little attention has been paid to the reliability of neural entrainment. This is despite the fact that the reliability of entrainment is of paramount importance both for understanding its contribution to auditory perception and for developing effective targeted interventions. Here, we showed that neural entrainment to FM sounds is reliable over time: we observed high inter-session correlations for metrics of entrainment strength, such as the spectral amplitude of neural activity at the stimulus frequency, and perhaps more importantly, the phase lag between stimulus and brain. The fact that entrainment signatures are stable over time (at least for FM sounds) confirms that rhythmic auditory stimulation is a valid tool for the external manipulation of narrow-band brain activity and as such might be useful for restoring or facilitating oscillatory brain dynamics (36, 37). Moreover, rhythmic auditory stimulation could be used to test the causal role of brain rhythms for different brain functions, either alone or in combination with other non-invasive stimulation techniques

(e.g., transcranial alternating current stimulation or transcranial magnetic stimulation); manipulating the relationship between the FM stimulus and the electrical stimulation could in principle be used as a direct manipulation between the electrical manipulation and brain activity.

Entrained delta and ongoing alpha activity influenced gap-detection performance

Here, we observed that trial-by-trial gap-detection performance was mainly predicted by pre-gap neural 2-Hz phase. Moreover, both delta and alpha phase predicted the observed FM-driven behavioral modulation, with their interaction providing stronger predictive power than either predictor considered alone. These results are in line with previous studies showing that the phase of neural oscillations prior to target occurrence predicts perception in different sensory domains (8, 38-41). For example, enhanced detection or faster reaction times have been reported for sensory stimuli presented at the optimal phase of delta (42), theta (43, 44), or alpha oscillations (39, 44). Our results also showed that, in a rhythmic listening context, gap-detection performance was not just a product of the entrained delta activity but also depended on non-entrained (ongoing) alpha activity. Comodulation of behavior by different frequencies have previously been reported (6). Critically, such comodulation was specific to the entrained frequencies, which led the authors to conclude that environmental rhythms reduce dimensionality of neural dynamics. Here we expand this view by showing that both entrained and ongoing brain oscillations could potentially comodulate behavioral performance. We interpret the contributions of both entrained and ongoing activity to perception as potentially reflecting an interplay between stimulus driven (sensory, bottom-up) and internally driven (top-down) processes. Although our paradigm was not designed to provide a time-resolved look at the interplay between delta and alpha oscillations, our results are consistent with alternating influences that might occur as a result of "lapses" of entrainment during which alpha oscillations might have a stronger effect on perception (10). It is possible that when entrainment is high, auditory perception is mostly modulated by the entrained activity, however, during entrainment lapses, perception is shaped by internal ongoing activity. Or conversely, participants could have adopted the strategy of trying to ignore the rhythm in order to perform the task, but as their alpha–attention system lapsed, they were forced into a rhythmic–entrainment mode.

Quite a bit of previous work has demonstrated the importance of alpha activity for "gating" near-threshold stimuli into awareness (8, 45), which goes beyond the idea of alpha activity indexing lapses of entrainment. However, the vast majority of these studies have been

conducted in the visual modality (46-48), and it has been argued that alpha activity does not contribute to near-threshold auditory perception (49). Thus, one open question relates to the precise role of alpha oscillations in shaping near-threshold auditory perception, in particular in a rhythmic auditory context. Answering this question requires respecting the observation of different alpha rhythms with different neural generators (e.g., (50-53)). Moreover, several studies linking alpha oscillations to attention have suggested that alpha oscillations could play at least two different roles: i.e., a facilitatory role where it can enhance target processing, or a suppressive role where alpha activity can suppress the processing of distractors (54-56). We hypothesize that the alpha-phase effect observed in our study is reflecting something like distractor suppression. While delta oscillatory activity entrained to the FM stimulus facilitates target processing at the optimal delta phase, we speculate that ongoing alpha activity might play a role in suppressing the distracting stimulus itself (complex noise), in an attempt to also maximize target detection. In this case, stimulus-induced sinusoidal behavioral modulation would be best predicted by the interaction of both mechanisms, in line with the current results.

Behavioral entrainment consistency depends on modulation rate

In the current study, we were surprised to observe that the preferred *stimulus* phase that yielded best gap-detection performance was consistent across participants. In previous work, preferred stimulus phase was uniformly distributed across participants, and this observation was critical to our argument that stimulus-driven behavioral modulation was the result of neural entrainment and not an artifact of stimulus acoustics (7). The primary difference between the current study and previous work was the FM rate, which was slower here (2 Hz) than in previous work (3 Hz). This led us to conduct a control experiment where we examined gap-detection performance for stimuli with FM rates varying between 1.5 and 3 Hz (in 0.5-Hz steps). We found that preferred phase became less consistent as FM rate increased (but only when a quadratic term was also included as a predictor).

Humans prefer to listen to and interact with auditory stimuli that are characterized by rhythmic structure in delta band, with a mode around 2 Hz (57-60). This rate overlaps with the modal periodicities of the human body (61), the most common period for the beat rate in Western music (61), and to common speech rates as quantified by inter-stress and inter-word levels in a variety of languages (62) and EMG data during speaking (63). We hypothesize that because the 2-Hz stimuli in the current study better aligned with preferred rate for humans, they may have been more consistently tracked by listeners compared to faster stimuli that were less likely to correspond to preferred rates and may have led to more phase slips / precession (64) and

therefore less consistency across participants. We note that there is evidence that favored rates for different types of stimuli may differ (e.g., speech vs. music (65); auditory vs. visual rhythms (66, 67)). Moreover, behavioral preferred rates change over the lifespan (60, 68), as does neural entrainment to different stimuli (FM/AM) (26). Human participants also differ in their output frequency e.g., when asked to talk or tap at a comfortable rate (69, 70), which is interpreted as individual differences in optimal frequencies for producing or processing incoming sensory information. The interaction between stimulus rate and individual differences in neural oscillator properties, including resonance frequency, is still a matter for empirical work. However, the results of our behavioral study, taken together with previous behavioral (60) and electrophysiological work (63, 64) suggest that neural entrainment is more successful and consistent when stimulus rates more closely match individual preferred rates.

Conclusion

Taken together, our results showed that FM stimuli entrained neural activity and sinusoidally modulated near-threshold target detection: both signatures of entrainment as well as its behavioral consequences were reliable across sessions. This demonstration is a critical prerequisite for research lines focused on targeted interventions for entrainment but has to our knowledge gone untested until now. Moreover, gap-detection performance was predicted by entrained neural delta phase, ongoing alpha phase and their interaction, suggesting that delta and alpha phase underpin different but potentially simultaneously active neural mechanisms and together shape perception.

Methods

Main experiment

Participants

Main study. Forty-one healthy participants took part in the study. Three participants were excluded from further analysis due to noisy EEG data (1 participant) and poor task performance (i.e., detection rate <0.25, 2 participants). Results presented in this manuscript include data from 38 participants (21 females, four left-handed, mean age: 26.03 with SD = 4.6 years old). Each participant took part in two sessions separated by 2–42 days (median: 7 days). All participants self-reported normal-hearing and normal or corrected-to-normal vision. All participants were either native German speakers (n=37) or spoke German with high proficiency (n=1). At the

time of the experiment no participant was taking medication for any neurological or psychiatric disorder.

Participants received financial compensation for their participation in the study. Written informed consent was obtained from all participants. The procedure was approved by the Ethics Council of the Max Planck Society and in accordance with the declaration of Helsinki.

Stimuli

Auditory stimuli were generated by MATLAB software at a sampling rate of 44,100 Hz. Stimuli were 20-s long complex tones frequency modulated in at a rate of 2 Hz and a center-to-peak depth of 67% (**Fig. 1**). The center frequency for the complex carrier signals was randomly chosen for each stimulus within the range of 1000-1400Hz. The complex carrier comprised 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, $\pi/4$, $\pi/2$, $3\pi/4$, π, $5\pi/4$, $3\pi/2$, $7\pi/4$) with the constraint that each trial would always start with a phase different to its predecessor. All stimuli were rms amplitude normalized. Three, four, or five silent gaps were inserted into each 20-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 18 equally spaced phase bins into which each single cycle of the frequency modulation was divided. No gaps were presented either in the first or the last second of the stimulus. A minimum of 1.5 s separated consecutive gaps.

Procedure

The experiment was conducted in an electrically shielded and acoustically isolated chamber and under normal-illumination conditions. Sound-level thresholds were determined for each participant according to the method of limits. All stimuli were then presented at 55 dB above the individual hearing threshold (55 dB sensation level, SL).

Gap duration was individually adjusted to detection threshold levels using an adaptive-tracking procedure comprising two interleaved staircases and a weighted up-down technique with custom weights. During this procedure, participants detected a gap within a 4-s sound. Except for the duration, the sound had the same characteristics as in the main experiment. The

descending staircase started with a gap duration of 150 ms and the ascending staircase started with a gap duration of 1 ms. If the participant detected the gap, gap duration was decreased by some percent (5% for 10 ms ≤ gaps ≤ 35 ms, 20% for 35 ms < gaps ≤ 70 ms, or 50% for 70 ms < gaps || gaps < 10 ms) in the following trial of the current staircase. On the contrary, if the participant did not detect the gap, gap duration was increased by some percent (following the same convention as before) of the current gap duration, in the following trial of the current staircase. Each staircase ended when four reversals occurred in a span of six trials. The mean final gap duration across the two staircases was chosen for presenting the gaps in the main task.

Before starting the main experiment, participants performed practice trials to make sure they understood the task. For the main experiment, EEG was recorded while listeners detected gaps embedded in the 20-s long FM stimuli. Listeners were instructed to respond as quickly as possible when they detected a gap via button-press. Overall, each listener heard 216 stimuli (27 per starting phase). The number of gaps per stimulus was counterbalanced (72 stimuli each included 3 gaps, 4 gaps, and 5 gaps) for a total of 864 gaps. For each of the 18 FM-phase bins, 48 gaps were presented. Including the EEG preparation, each experimental session lasted about 3 hours.

Data Acquisition and Analysis

Behavioral data. Behavioral data were recorded online by MATLAB 2017a (MathWorks) in combination with Psychtoolbox. Sounds were presented at a rate of 44.1kHz, via an external soundcard (RME Fireface UCX 36-channel, USB 2.0 & FireWire 400 audio interface) using ASIO drivers. Participants listened to the sounds via over-ear headphones (Beyerdynamic DT-770 Pro 80 Ohms, Closed-back Circumaural Dynamic Diffuse field equalization Impedance: 80 Ohm SPL: 96 dB Frequency range: 5–35,000 Hz). Button presses were collected using a Cedrus response pad (RB-740). Hits were defined as button-press responses that occurred no earlier than 100 ms and no later than 1.5 s after the occurrence of a gap. Hit rates and RTs were calculated separately for each of the 18 FM-phase bins. To estimate the FM-induced sinusoidal modulation of gap detection behavior, a cosine function was fitted to hit rates as a function of FM-phase for each participant and each session. From the fitted function, the amplitude parameter quantifies the strength of behavioral modulation by 2-Hz FM phase, while the phase parameter indexes the FM-stimulus–brain lag. Significance of the sinusoidal modulation was tested using a permutation approach, whereby 1000 surrogate datasets were created for each participant and session by shuffling the single-gap accuracy values (0,1) with respect to their stimulus-condition labels. Cosine functions were also fitted to the surrogated datasets. Each

participant's amplitude parameter was converted to z-score using the mean and standard deviation of the individual participant surrogate datasets. Gap detection was considered to be sinusoidally modulated for each participant if the z-score of the fitted amplitude parameter exceeded z = 3.2 (i.e., $p<0.05$, *Bonferroni* corrected for 76 comparisons -i.e., 38 subjects x two sessions). Preferred FM-phase was defined as the instantaneous phase of the fitted function with the highest hit rate.

In order to test the effect of stimulus characteristics and time on gap detection, logistic regression models were fitted individually for each participant and session using the MATLAB function '*fitglm*', using binomial distribution and logit as the link function. Collinearity between regressors was assessed using the MATLAB function "*collintest*". Different models were tested evaluating whether gap detection could be predicted as a function of 1) FM-phase, 2) stimulus center carrier frequency (1000-1400Hz), 3) stimulus center carrier at gap onset, 4) global time (1-864, indicating the gap position within the whole experiment), 5) local time (1-12 time bins, indicating gap location within the 20-s stimulus) and the interactions of 6) FM-phase by stimulus center carrier frequency, 7) FM-phase by global time and 8) FM-phase by local time. Five different models were fitted, each including a different combination of regressors (Table S2) and the best model was selected using the Akeike's information criterium corrected for small samples (AICc). Before running the models, all predictors involving circular data (phase angles) were linearized by calculating their sine and cosine. The overall beta estimate for the final predictor (b) was then calculated by combining the beta estimates of the sine (sinb) and cosine (cosb) predictors:

$$b = \sqrt{(sinb^2 + cosb^2)}$$

To test the effect of each predictor at the group level, 1000 surrogate datasets were created for each participant and session by shuffling the single-gap accuracy values (0,1) while keeping the stimulus conditions the same. The same regression models were fitted to the surrogate datasets and the mean beta estimate for each regressor was taken as the random distribution mean. One-sample *t-tests* against the random distribution mean were conducted to assess the significance of each regressor at the group level, separately for each session. Dependent sample *t-tests* were conducted to compare between sessions. *P-values* were corrected for multiple comparisons using *Bonferroni* method.

Control study. Sixteen participants took part in the control experiment (12 females, mean age 27.1 (SD = 5)). Ten participants had also been recruited for the main study, they received

financial compensation and signed written informed consent. The other six participants were colleagues in our research group (including one author-YCC) and participated voluntarily without compensation).

Unless otherwise specified, stimuli and procedures were defined as for the main experiment. In contrast to the main experiment, gap thresholds were not individually defined but gap duration was fixed at 16 ms (mean threshold in session 2 in the main experiment) for all participants. FM-stimuli were created as for the main experiment but modulated at four different frequencies, i.e., 1.5, 2, 2.5, and 3 Hz. Gaps could be presented at 15 different bins of the FM cycle. Subjects heard 224 stimuli (56 per FM) for a total of 932 gaps (233 per FM rate). Stimuli were presented in 8 different blocks (2 per FM rate) of 28 stimuli each. Each block comprised only one FM rate but the FM order was randomized within session and across participants. The number of gaps per stimulus (3, 4, or 5) was randomized. For each of the 15 FM-phase bins*FM rate combination, 14-16 gaps were presented.

Hits were defined as button-press responses that occurred no earlier than 100 ms and no later than 1.5 s after the occurrence of a gap. Hit rates were calculated separately for each of the 15 FM-phase bins and each FM rate. To estimate the FM-induced sinusoidal modulation of gap detection behavior, a cosine function was fitted to hit rates as a function of FM-phase for each participant. For each participant and FM rate, the mean hit rate (i.e., fitted intercept), the fitted amplitude parameter and the preferred phase (same definition as in the main experiment) were estimated. The effect of FM rate on hit rates and amplitude parameters was tested using a one-way Analysis of Variance (ANOVA). To test whether individual preferred phases were more randomly distributed with increasing FM rate, resultant vector lengths were computed for each FM rate using the individual preferred phases. Two linear models were fitted to the resultant vector length by FM rate data. The first one included intercept and linear (FM rate) terms while to second also included a quadratic term (FM ^2). Both models were fitted to the data using the Matlab function *"fitglm"* and statistical significance was estimated using permutation tests. For comparison, the test distribution was created by computing 1000 resultant vector lengths calculated with the individual preferred phases while swapping the FM rate information at the individual level.

Electroencephalogram data. The EEG was recorded with an actiCAP active electrode system in combination with Brainamp DC amplifiers (Brain Products GmbH). The electrode system included 64 Ag–AgCl electrodes mounted on a standard cap, actiCAP 64Ch Standard-2 (Brain Products GmbH). Signals were recorded continuously with a passband of 0.1 to 1000 Hz and

digitized at a sampling rate of 1000 Hz. For recording, the reference electrode was placed over FCz and the ground electrode over AFz. For better stimulus marking, in addition to standard EEG triggers from the LPT port, stimulus markers were also sent via soundcard and collected in the EEG using a Stimtrak (Brain Products GmbH). Electrode resistance was kept under 20 kΩ. All EEG data were analyzed offline by using Fieldtrip software (www.ru.nl/fcdonders/fieldtrip; version 20200130), and custom MATLAB scripts.

Two different preprocessing pipelines were implemented. One was tailored to assess entrainment characteristics and reliability and focused on the complete 20-s stimulus periods; the second pipeline was tailored to test the effect of pre-target (pre-gap) activity on gap detection, and focused on the periods around the gap's occurrence. In the first preprocessing pipeline, the continuous EEG data were high-pass filtered at 0.6 Hz. Filtered data were then epoched into 21.5-s trials (1 s before stimulus onset and 0.5 s after stimulus offset). The trial data were low-pass filtered at 80 Hz and the 50 Hz line noise was removed using discrete Fourier transform (dft) with spectrum interpolation as implemented in Fieldtrip. Data were re-referenced to the average reference. Extreme artifacts were removed based on visual inspection. Noisy electrodes were then interpolated (1 electrode in 3 participants and 2 electrodes in one participant). Eye-blinks, muscle, heartbeat, and remaining line noise or faulty contact artifacts were removed using ICA. Next, data were low pass filtered to 30 Hz and trials for which the range exceeded 200 uV were automatically removed. If more than 30% of the trials had to be removed because of artifacts, the participant was removed for further analysis (1 participant). Preprocessed data were resampled to 500 Hz.

The second preprocessing pipeline included the same steps excepting the initial high-pass filter. To maximize comparability with the first pipeline, all the same trials and ICA components that were identified based on the first pipeline were removed in the second pipeline. After all preprocessing steps and before resampling, 3 s long trials were defined around each gap onset (i.e., 1.5 s before and 1.5 s after gap onset). Trials exceeding a range of 200 uV were excluded and data were resampled to 500 Hz.

Frequency and time-frequency analysis of full-stimulus periods. Full-stimulus epochs were analyzed in the frequency and time-frequency domains to examine brain responses entrained by the 2-Hz stimulation. Since the starting phase of the FM stimulus was randomized from trial to trial, before conducting frequency-domain analyses, single-trial brain responses were realigned so that the FM stimulus phases would be perfectly phase-locked across trials after the realignment. A fast Fourier transform (FFT) was performed on the trial-averaged time-domain

data, after multiplication with a Hann window. Evoked amplitude in each frequency band was calculated as the absolute value of the complex output, while the phase angle of the complex FFT output at 2 Hz provided an estimate of stimulus-brain phase lag. An FFT was also applied on each single trial, and the resulting single-trial amplitude spectra were averaged over trials as an indicator of total amplitude of neural activity that was not necessarily phase-locked to the stimulus. Inter-trial phase coherence (ITPC) was calculated as the resultant vector length of phase angles from the complex FFT output across trials separately for each frequency and electrode. In addition, the single-trial time-domain data were submitted to a time–frequency analysis by using the Fieldtrip-implemented version of the Wavelet approach using Fourier output. Here, wavelet size varied with frequency linearly from three to seven cycles over the range from 1 to 15 Hz. The resulting complex values were used to estimate time-resolved ITPC for each channel separately.

To statistically test spectral amplitudes and ITPC at frequencies of interest (2 Hz, 4 Hz, alpha), nonparametric Wilcoxon signed rank test (in session 1: for 4 Hz ITPC; in session 2: for 2 Hz amplitude and ITPC and for alpha amplitude) or the parametric equivalent paired-samples *t*-tests (all other comparisons) were conducted, based on satisfaction of normality assumptions; whether to use parametric of non-parametric tests for comparisons was decided based whether the data was normally distributed or not according to the Lilliefors test implementation in MATLAB (function '*lillietest'*). For each condition, participant, and session, data were averaged over all channels (and over time for time-resolved ITPC) and amplitudes/ITPC of the two target frequencies (2 Hz and 4 Hz) were then tested against the average amplitude/ITPC of the neighboring ±8 frequency bins (0.16Hz) similar to (7, 27). In the case of alpha amplitude, data were averaged across all bins including frequencies between 7-12 Hz and were tested against the average amplitude of the neighboring ±100 frequency bins (2 Hz).

Based on the topography of the 2/4 Hz and alpha (7-12 Hz) amplitude spectra, further analyses involving the extraction of phase and/or amplitude values were done in a cluster of electrodes including *F3, Fz, F4, FC1, FC2, C3, Cz, C4, F1, F2, FC3, FC4, C1, C2* for 2/4 Hz activity and *P8, P6, P4, P2, Pz, P1, P3, P5, P7, PO9, PO10, PO8, PO4, POz, PO3, PO7, O1, Oz, O2* for alpha activity.

Pre-gap activity. Before analysis of pre-gap activity, single-trial time-domain data around the gap period (1.5 s before and 1.5 s after) were detrended (using linear regression). It is possible that the smearing of the evoked response back into pre-stimulus period by wavelet convolution could produce spurious pre-stimulus phase effects. To minimize this, gap-evoked responses

were removed from the post-stimulus period by multiplication with half of a Hann window that ranged between 0 and 50 ms after gap onset and was zero thereafter (6). Next, we applied two different analysis approaches to quantify neural phase and neural amplitude in the pre-gap time window. To estimate neural amplitude, the time-domain data were submitted to a wavelet convolution using Fourier output as implemented in Fieldtrip. Wavelet size varied with frequency linearly from three to seven cycles over the range from 1 to 15 Hz with 10 ms temporal resolution. Alpha amplitude as well as 2-Hz amplitude were averaged within the 100-ms time window preceding gap onset.

For extracting the pre-gap instantaneous 2 Hz and alpha phase, single-trial data were first band-pass filtered using a Butterworth filtered as implemented in Fieldtrip (1.5.-2.5 Hz for 2 Hz activity and 7-12 Hz for alpha activity). Filtered data were subjected to a Hilbert transform and the phase angle was computed. Time windows for extracting pre-gap 2 Hz and alpha phases were adjusted to include 1/5 of a cycle of the relevant frequency, i.e., 100 ms preceding gap onset for 2 Hz phase and 22 ms for alpha (assuming center frequency 9 Hz) phase.

The influence of the instantaneous 2-Hz phase and amplitude, alpha phase and amplitude, and 2-Hz–alpha phase amplitude coupling (PAC) in the pre-gap period on gap detection was evaluated using logistic regression models at the single-participant level, similar to the behavioral analysis using the MATLAB function *'fitglm'*, specifying the distribution as binomial and the link function as logit. Collinearity between regressors was assessed using the MATLAB function "collintest". Six different logistic regression models were fitted to the data including a different combination of the regressors mentioned above and the best model was chosen based on the AICc value (**Table S1**). To test the effect of each predictor at the group level, 1000 surrogate datasets were created for each participant and session by shuffling the single-gap accuracy values (0,1) across trials while keeping the regressors the same. The same regression models where fitted to the surrogate datasets and the mean beta estimate for each regressor was taken as the random distribution mean. One-sample *t-tests* against the random distribution mean were conducted to assess the significance of each regressor at the group level, separated by session. Dependent sample *t-tests* were conducted to compare between sessions. *P-values* were corrected for multiple comparisons using *Bonferroni* method.

To investigate optimal 2-Hz phase angle for gap detection, trials were sorted according to the pre-gap 2-Hz phase angles (grouped in 18 equally spaced phase bins) and hit rates were calculated for each bin. Similarly, to estimate optimal alpha phase angle for gap detection, trials were sorted according to the pre-gap alpha phase angles (grouped in 18 equally spaced phase

bins) and hit rates were calculated for each bin. Phase lag was estimated based on a cosine fit to hit rates binned by either delta or alpha phase. Separately for delta and alpha, best neural phase was defined as the phase with highest detection rate as estimated from the cosine fit.

Predicting FM-stimulus driven sinusoidal behavioral modulation from neural phase effects. Three detection probabilities were assigned to each gap: 1) the hit rate calculated for the pre-gap neural 2-Hz phase bin into which the gap fell, 2) the hit rate calculated for the pre-gap neural alpha phase bin into which the gap fell, and 3) the interaction of the two (their product, **Fig. 6a**). Then, gaps were sorted and binned based on their FM-*stimulus* phase bin (18 bins), and three predicted functions were calculated by averaging the predicted hit rates across gaps in each bin. Pearson correlation coefficients were computed between the true individual observed hit rates by FM-stimulus phase bin and each function predicted in the same session (within session prediction) or in the opposite session (inter-session prediction). For statistical analysis, 1000 surrogate datasets per subject and session were created by shuffling the bin labels in the true hit-rate by FM-stimulus profiles. As with the true profiles, each predicted function was correlated with the surrogate hit-rate by phase bin profile from the same (within session prediction) and the opposite session (inter-session prediction). Before further analyses, all correlation coefficients were z-scored using Fisher's r-to-z transformation. For each predicted function, statistical significance was estimated using one-sample t-test (comparing each predictor vs. the mean Fisher's z-score from the surrogate data analysis). Significant difference between sessions and predictors were estimated using paired-samples t-tests. All p-values were corrected using *Bonferroni* method.

Questionnaires

To evaluate musical skills, all participants from the main experiment complete the Goldsmiths Musical Sophistication Index (Gold-MSI) (30). Scores were computed using the documents and templates provided in https://www.gold.ac.uk/music-mind-brain/gold-msi/download/. Accordingly, individual scores were extracted indexing five main factors (i.e., active engagement, perceptual abilities, musical training, emotions, singing abilities) and the general sophistication index.

Statistical Analysis

Prior to any statistical analysis, data normality was tested using the Lilliefors normality test in MATLAB. Parametric or equivalent non-parametric tests were then chosen to test for

significant differences or correlations between variables. The test showed a significant deviation from normality for the false alarms and final gap size, accordingly non parametric tests were chosen to test for correlation (Spearman) and significant difference between sessions (Wilcoxon signed rank test). Unless otherwise specified in the text, all other correlation analyses were done using Pearson correlation coefficient for linear variables and circular-circular or circular-linear correlation when circular data were involved. Significant difference between sessions were tested using one-sample or dependent-samples *t-test* when normality assumptions were satisfied. For performing statistical comparisons on correlation coefficients, coefficients were always z-scored using the Fisher's r-to-z transformation method. Unless otherwise specified, significant p-values were corrected using *Bonferroni* method.

Acknowledgements

We thank Nicole Huizinga for helping with the data collection. We thank Hanna Kadel, Dominik Thiele, and Johannes Messerschmidt for technical support and for helping with the EEG preparation and participants recruitment. We also thank Cornelius Abel for technical support. This work was supported by an ERC Starting Grant (BRAINSYNC) awarded to MJH.

Author contributions

YCC designed the study, collected and analyzed the data, and wrote the manuscript. MJH designed the study, supervised data analysis and wrote the manuscript. Both authors agree with the final version of the manuscript.

Competing interests

The authors declare no competing interest.

References

1. Obleser J, Kayser C. Neural Entrainment and Attentional Selection in the Listening Brain. Trends Cogn Sci. 2019;23(11):913-26.
2. Lakatos P, Karmos G, Mehta AD, Ulbert I, Schroeder CE. Entrainment of neuronal oscillations as a mechanism of attentional selection. Science. 2008;320(5872):110-3.
3. Giraud AL, Poeppel D. Cortical oscillations and speech processing: emerging computational principles and operations. Nat Neurosci. 2012;15(4):511-7.
4. Lakatos P, Shah AS, Knuth KH, Ulbert I, Karmos G, Schroeder CE. An oscillatory hierarchy controlling neuronal excitability and stimulus processing in the auditory cortex. J Neurophysiol. 2005;94(3):1904-11.
5. Buzsaki G, Draguhn A. Neuronal oscillations in cortical networks. Science. 2004;304(5679):1926-9.

6. Henry MJ, Herrmann B, Obleser J. Entrained neural oscillations in multiple frequency bands comodulate behavior. Proc Natl Acad Sci U S A. 2014;111(41):14935-40.
7. Henry MJ, Obleser J. Frequency modulation entrains slow neural oscillations and optimizes human listening behavior. Proc Natl Acad Sci U S A. 2012;109(49):20095-100.
8. Mathewson KE, Prudhomme C, Fabiani M, Beck DM, Lleras A, Gratton G. Making waves in the stream of consciousness: entraining oscillations in EEG alpha and fluctuations in visual awareness with rhythmic visual stimulation. J Cogn Neurosci. 2012;24(12):2321-33.
9. Ten Oever S, Schroeder CE, Poeppel D, van Atteveldt N, Mehta AD, Megevand P, et al. Low-Frequency Cortical Oscillations Entrain to Subthreshold Rhythmic Auditory Stimuli. J Neurosci. 2017;37(19):4903-12.
10. Lakatos P, Barczak A, Neymotin SA, McGinnis T, Ross D, Javitt DC, et al. Global dynamics of selective attention and its lapses in primary auditory cortex. Nat Neurosci. 2016;19(12):1707-17.
11. Garcia-Rosales F, Beetz MJ, Cabral-Calderin Y, Kossl M, Hechavarria JC. Neuronal coding of multiscale temporal features in communication sequences within the bat auditory cortex. Commun Biol. 2018;1:200.
12. Doelling KB, Arnal LH, Ghitza O, Poeppel D. Acoustic landmarks drive delta-theta oscillations to enable speech comprehension by facilitating perceptual parsing. Neuroimage. 2014;85 Pt 2:761-8.
13. Doelling KB, Poeppel D. Cortical entrainment to music and its modulation by expertise. Proc Natl Acad Sci U S A. 2015;112(45):E6233-42.
14. Zion Golumbic EM, Ding N, Bickel S, Lakatos P, Schevon CA, McKhann GM, et al. Mechanisms underlying selective neuronal tracking of attended speech at a "cocktail party". Neuron. 2013;77(5):980-91.
15. Horton C, Srinivasan R, D'Zmura M. Envelope responses in single-trial EEG indicate attended speaker in a 'cocktail party'. J Neural Eng. 2014;11(4):046015.
16. Horton C, D'Zmura M, Srinivasan R. Suppression of competing speech through entrainment of cortical oscillations. J Neurophysiol. 2013;109(12):3082-93.
17. Brodbeck C, Jiao A, Hong LE, Simon JZ. Neural speech restoration at the cocktail party: Auditory cortex recovers masked speech of both attended and ignored speakers. PLoS Biol. 2020;18(10):e3000883.
18. Zoefel B, Archer-Boyd A, Davis MH. Phase Entrainment of Brain Oscillations Causally Modulates Neural Responses to Intelligible Speech. Curr Biol. 2018;28(3):401-8 e5.
19. Riecke L, Formisano E, Sorger B, Baskent D, Gaudrain E. Neural Entrainment to Speech Modulates Speech Intelligibility. Curr Biol. 2018;28(2):161-9 e5.
20. Riecke L, Sack AT, Schroeder CE. Endogenous Delta/Theta Sound-Brain Phase Entrainment Accelerates the Buildup of Auditory Streaming. Curr Biol. 2015;25(24):3196-201.
21. Wilsch A, Neuling T, Obleser J, Herrmann CS. Transcranial alternating current stimulation with speech envelopes modulates speech comprehension. Neuroimage. 2018;172:766-74.
22. Cabral-Calderin Y, Wilke M. Probing the Link Between Perception and Oscillations: Lessons from Transcranial Alternating Current Stimulation. Neuroscientist. 2020;26(1):57-73.
23. Erkens J, Schulte M, Vormann M, Herrmann CS. Lacking Effects of Envelope Transcranial Alternating Current Stimulation Indicate the Need to Revise Envelope Transcranial Alternating Current Stimulation Methods. Neurosci Insights. 2020;15:2633105520936623.
24. Henry MJ, Herrmann B. A precluding role of low-frequency oscillations for auditory perception in a continuous processing mode. J Neurosci. 2012;32(49):17525-7.
25. Schroeder CE, Lakatos P. Low-frequency neuronal oscillations as instruments of sensory selection. Trends Neurosci. 2009;32(1):9-18.

26. Henry MJ, Herrmann B, Kunke D, Obleser J. Aging affects the balance of neural entrainment and top-down neural modulation in the listening brain. Nat Commun. 2017;8:15801.
27. Bauer AR, Bleichner MG, Jaeger M, Thorne JD, Debener S. Dynamic phase alignment of ongoing auditory cortex oscillations. Neuroimage. 2018;167:396-407.
28. Ding N, Simon JZ. Power and phase properties of oscillatory neural responses in the presence of background activity. J Comput Neurosci. 2013;34(2):337-43.
29. Zoefel B, Davis MH, Valente G, Riecke L. How to test for phasic modulation of neural and behavioural responses. Neuroimage. 2019;202:116175.
30. Mullensiefen D, Gingras B, Musil J, Stewart L. The musicality of non-musicians: an index for assessing musical sophistication in the general population. PLoS One. 2014;9(2):e89642.
31. Meyer L, Sun Y, Martin AE. Synchronous, but not entrained: exogenous and endogenous cortical rhythms of speech and language processing. Lang Cogn Neurosci. 2020;35(9):1089-99.
32. Rimmele JM, Morillon B, Poeppel D, Arnal LH. Proactive Sensing of Periodic and Aperiodic Auditory Patterns. Trends Cogn Sci. 2018;22(10):870-82.
33. Lakatos P, Gross J, Thut G. A New Unifying Account of the Roles of Neuronal Entrainment. Curr Biol. 2019;29(18):R890-R905.
34. Sameiro-Barbosa CM, Geiser E. Sensory Entrainment Mechanisms in Auditory Perception: Neural Synchronization Cortico-Striatal Activation. Front Neurosci. 2016;10:361.
35. Ng BS, Schroeder T, Kayser C. A precluding but not ensuring role of entrained low-frequency oscillations for auditory perception. J Neurosci. 2012;32(35):12268-76.
36. Harrington MO, Ashton JE, Ngo HV, Cairney SA. Phase-locked Auditory Stimulation of Theta Oscillations during Rapid Eye Movement Sleep. Sleep. 2020.
37. Calabro RS, Naro A, Filoni S, Pullia M, Billeri L, Tomasello P, et al. Walking to your right music: a randomized controlled trial on the novel use of treadmill plus music in Parkinson's disease. J Neuroeng Rehabil. 2019;16(1):68.
38. Hanslmayr S, Aslan A, Staudigl T, Klimesch W, Herrmann CS, Bauml KH. Prestimulus oscillations predict visual perception performance between and within subjects. Neuroimage. 2007;37(4):1465-73.
39. Vanrullen R, Busch NA, Drewes J, Dubois J. Ongoing EEG Phase as a Trial-by-Trial Predictor of Perceptual and Attentional Variability. Front Psychol. 2011;2:60.
40. Wostmann M, Waschke L, Obleser J. Prestimulus neural alpha power predicts confidence in discriminating identical auditory stimuli. Eur J Neurosci. 2019;49(1):94-105.
41. Weisz N, Wuhle A, Monittola G, Demarchi G, Frey J, Popov T, et al. Prestimulus oscillatory power and connectivity patterns predispose conscious somatosensory perception. Proc Natl Acad Sci U S A. 2014;111(4):E417-25.
42. Stefanics G, Hangya B, Hernadi I, Winkler I, Lakatos P, Ulbert I. Phase entrainment of human delta oscillations can mediate the effects of expectation on reaction speed. J Neurosci. 2010;30(41):13578-85.
43. Busch NA, VanRullen R. Spontaneous EEG oscillations reveal periodic sampling of visual attention. Proc Natl Acad Sci U S A. 2010;107(37):16048-53.
44. Busch NA, Dubois J, VanRullen R. The phase of ongoing EEG oscillations predicts visual perception. J Neurosci. 2009;29(24):7869-76.
45. Mathewson KE, Gratton G, Fabiani M, Beck DM, Ro T. To see or not to see: prestimulus alpha phase predicts visual awareness. J Neurosci. 2009;29(9):2725-32.
46. Fodor Z, Marosi C, Tombor L, Csukly G. Salient distractors open the door of perception: alpha desynchronization marks sensory gating in a working memory task. Sci Rep. 2020;10(1):19179.

47. Spaak E, de Lange FP, Jensen O. Local entrainment of alpha oscillations by visual stimuli causes cyclic modulation of perception. J Neurosci. 2014;34(10):3536-44.
48. Hutchinson BT, Pammer K, Jack B. Pre-stimulus alpha predicts inattentional blindness. Conscious Cogn. 2020;87:103034.
49. VanRullen R, Zoefel B, Ilhan B. On the cyclic nature of perception in vision versus audition. Philos Trans R Soc Lond B Biol Sci. 2014;369(1641):20130214.
50. Keitel C, Keitel A, Benwell CSY, Daube C, Thut G, Gross J. Stimulus-Driven Brain Rhythms within the Alpha Band: The Attentional-Modulation Conundrum. J Neurosci. 2019;39(16):3119-29.
51. Isaichev SA, Derevyankin VT, Koptelov Yu M, Sokolov EN. Rhythmic alpha-activity generators in the human EEG. Neurosci Behav Physiol. 2001;31(1):49-53.
52. Weisz N, Hartmann T, Muller N, Lorenz I, Obleser J. Alpha rhythms in audition: cognitive and clinical perspectives. Front Psychol. 2011;2:73.
53. Billig AJ, Herrmann B, Rhone AE, Gander PE, Nourski KV, Snoad BF, et al. A Sound-Sensitive Source of Alpha Oscillations in Human Non-Primary Auditory Cortex. J Neurosci. 2019;39(44):8679-89.
54. Schneider D, Herbst SK, Klatt L, Wöstmann M. Target Enhancement or Distractor Suppression? Functionally Distinct Alpha Oscillators form the Basis of Attention. PsyArXiv. 2020.
55. Bonnefond M, Jensen O. Alpha oscillations serve to protect working memory maintenance against anticipated distracters. Curr Biol. 2012;22(20):1969-74.
56. Wostmann M, Alavash M, Obleser J. Alpha Oscillations in the Human Brain Implement Distractor Suppression Independent of Target Selection. J Neurosci. 2019;39(49):9797-805.
57. Woodrow H. Chapter: Time perception. Handbook of experimental psychology. Oxford, England: Wiley; England; 1951. p. 1224-36.
58. Fraisse P. Rhythm and Tempo. In: Deutsch D, editor. Psychology of Music: Academic Press; 1982. p. 149-80.
59. van Noorden L, Moelants D. Resonance in the Perception of Musical Pulse. Journal of New Music Research. 1999;28(1):43-66.
60. McAuley JD, Jones MR, Holub S, Johnston HM, Miller NS. The time of our lives: life span development of timing and event tracking. J Exp Psychol Gen. 2006;135(3):348-67.
61. MacDougall HG, Moore ST. Marching to the beat of the same drummer: the spontaneous tempo of human locomotion. J Appl Physiol (1985). 2005;99(3):1164-73.
62. Allen GD. Speech rhythm: its relation to performance universals and articulatory timing. Journal of Phonetics. 1975;3(2):75-86.
63. Ruspantini I, Saarinen T, Belardinelli P, Jalava A, Parviainen T, Kujala J, et al. Corticomuscular coherence is tuned to the spontaneous rhythmicity of speech at 2-3 Hz. J Neurosci. 2012;32(11):3786-90.
64. Notbohm A, Kurths J, Herrmann CS. Modification of Brain Oscillations via Rhythmic Light Stimulation Provides Evidence for Entrainment but Not for Superposition of Event-Related Responses. Front Hum Neurosci. 2016;10.
65. Ding N, Patel AD, Chen L, Butler H, Luo C, Poeppel D. Temporal modulations in speech and music. Neurosci Biobehav Rev. 2017;81(Pt B):181-7.
66. Patel AD, Iversen JR, Chen Y, Repp BH. The influence of metricality and modality on synchronization with a beat. Exp Brain Res. 2005;163(2):226-38.
67. McAuley JD, Henry MJ. Modality effects in rhythm processing: Auditory encoding of visual rhythms is neither obligatory nor automatic. Atten Percept Psychophys. 2010;72(5):1377-89.
68. Drake C, Jones MR, Baruch C. The development of rhythmic attending in auditory sequences: attunement, referent period, focal attending. Cognition. 2000;77(3):251-88.

69. Collyer CE, Broadbent HA, Church RM. Preferred rates of repetitive tapping and categorical time production. Percept Psychophys. 1994;55(4):443-53.
70. Assaneo MF, Rimmele JM, Sanz Perl Y, Poeppel D. Speaking rhythmically can shape hearing. Nat Hum Behav. 2020.