Opposing neural processing modes alternate rhythmically during sustained auditory attention

Florian H. Kasten[1,2,*], Quentin Busson[2] and Benedikt Zoefel[1,2,*]

[1]Centre de Recherche Cerveau & Cognition, CNRS, Toulouse, France

[2]Université Toulouse III Paul Sabatier, Toulouse, France

*Corresponding authors:

Benedikt Zoefel, Florian H. Kasten

Centre de Recherche Cerveau & Cognition, CNRS, Toulouse, France

benedikt.zoefel@cnrs.fr

kasten.florian@cnrs.fr

Keywords: brain oscillations, entrainment, auditory system, sustained attention, speech perception

(Abstract: 235 words, Significance statement: 105 words, Manuscript: 2579 words, Methods: 1902 words, Number of Figures: 3)

Abstract

When confronted with continuous tasks, humans show spontaneous fluctuations in performance, putatively caused by varying attentional resources allocated to process external information. If neural resources are used to process other, presumably “internal” information, sensory input can be missed and explain an apparent dichotomy of “internal” versus “external” attention.

Each of these opposing attentional modes might have their own distinct neural signature. For instance, α-oscillations (~10-Hz) have been linked to a suppression of sensory information and might

therefore reflect a state of internal attention. Indeed, in auditory cortex of macaque monkeys, periods with strong α-oscillations and reduced responses to sounds alternate with periods in which the opposite pattern occurs. In the current study, we extracted neural signatures of internal and external attention in human electroencephalography (EEG) - α-oscillations and neural entrainment, respectively. We then tested whether they exhibit structured fluctuations over time, when listeners attended to an ecologically relevant stimulus like speech, and completed a task that required full and continuous attention.

Results showed an antagonistic relation between spontaneous α-oscillations and neural activity synchronized to speech. These opposing neural modes underwent slow, periodic fluctuations around ~0.07 Hz that were strikingly similar to those observed in non-human primates, and related to the successful detection of auditory targets. Our study might have tapped into a general attentional mechanism that is conserved across species and has important implications for situations in which sustained attention to sensory information is critical.

Significance

Understanding fluctuations in sustained attention is crucial for a wide variety of everyday activities, e.g., in traffic, education or safety monitoring. Lapses in attention can negatively affect outcomes in these tasks, sometimes with severe consequences. Here we demonstrate that auditory sustained attention fluctuates between an external mode favoring the processing of sensory information and an internal mode, where sensory information is ignored. These attentional modes have their own neural signatures and alternate with an inherent rhythmicity that can be traced in the human electroencephalogram. Our results may be leveraged to predict attentional fluctuations, optimize critical system designs or tailor interventional approaches to improve sustained attention.

1 Introduction

The ability to sustain attention is crucial for many activities of everyday life, yet it is surprisingly difficult to achieve[1,2]. Lapses in attention are common even in healthy populations and can have negative downstream effects on cognition[3] and lead to human error, sometimes with major consequences[4–6]. A variety of neurological and psychiatric disorders are characterized or accompanied by a decreased ability to maintain sustained attention[7,8]. Understanding the neural processes that give rise to dynamic fluctuations in attention is therefore critical.

It has been proposed that such fluctuations arise due varying amounts of attentional resources allocated to the processing of external information. When sensory input is ignored, neural resources might be available for other processes unrelated to external information (such as memory consolidation or internal thought), leading to a dichotomy of “internal” versus “external” attention[9,10]. Each of these opposing attentional modes might have its own signature, including specific patterns of neural connectivity[1,2] and neural oscillations[10–12]. Prominently, α-oscillations (~10-Hz) have been linked to suppression of sensory information[13,14] and attentional (de-)selection[15–18], and might therefore correspond to a state in which input is prone to be ignored.

Previous work in the auditory cortex of macaque monkeys supported the notion that internal attention is characterized by strong α-oscillations and reduced sensitivity to external information[10]. When subjects listened to rhythmic tone sequences, neural activity synchronized to the stimulus rhythm, an effect often termed neural entrainment[19–21]. At certain times, however, neural entrainment was reduced and α-oscillations dominated processing in auditory cortex. When α-oscillations prevailed, they rhythmically modulated neuronal firing and reduced neural and behavioral responses to stimulus input. Fundamental for our study, these bouts of α-oscillations (i.e. internal attention) occurred regularly and alternated with periods of strong entrainment to sound (i.e. external attention), at an inherent rhythm of ~0.06 Hz (i.e., ~16-sec)[10].

The identification of rhythmicity in attentional states and their neural counterparts has important implications for future research. It could be leveraged in the design of critical systems technology, educational environments or in the design of interventional approaches for situations where sustained attention is critical. However, it remained elusive if humans possess equivalent neural

signatures of attentional modes, and whether they exhibit any temporal regularity. It also remained unclear if regular attentional lapses occur during the processing of ecologically relevant stimuli. In particular, speech requires integration of information over time to be optimally perceived and therefore needs full and continuous attention[22].

We recorded electroencephalographic (EEG) data in humans and tested for regular fluctuations in attentional modes while they paid sustained attention to rhythmic speech sounds. We hypothesized that neural entrainment to speech and spontaneous α-oscillations show rhythmic fluctuations at ultra-slow frequencies (0.02-Hz - 0.2-Hz) and that these fluctuations show an antagonistic relationship (i.e. are coupled in anti-phase).

2 Results

2.1 Overview

We recorded participants' EEG while they listened to 5-min streams of rhythmic, monosyllabic, French words presented at a rate of 3-Hz (**Fig. 1a**). Depending on the experimental block, participants were instructed to keep their eyes-open or closed, respectively. They were asked to identify words that were presented off-rhythm (i.e., shifted by 80-ms relative to the 3-Hz rhythm). On average, participants detected 41.22% (± *SD*: 13.80) of targets during the eyes-open and 42.96% (± *SD*: 15.56) during eyes-closed (**Fig. 1b**) conditions. The proportion of false alarms was low relative to the large number of non-target words (eyes-open: 0.91% ± *SD*: 1.06, eyes-closed: 0.83% ± *SD*: 1.00, **Fig. 1c**). There was no difference in hits or false alarms between eyes-closed and eyes-open conditions (dependent samples t-test, hits: t_{22} = -0.94, p = .35, FA: t_{22} = 1.11 p = .27), nor was there a difference in reaction times (t_{22} = 1.62, p = .12: $M_{eyes\text{-}open}$ = 777-ms ± *SD*: 122; $M_{eyes\text{-}closed}$ = 738-ms ± *SD*: 96, **Fig. 1d**).

We used standard spectral analysis methods to extract spontaneous α-oscillations, and inter-trial coherence (ITC)[10,23] at 3-Hz to quantify auditory entrainment (see Materials and Methods). Both showed EEG topographies consistent with the literature[23,24] (**Fig 1e,f**). α-oscillations showed a dominant occipito-parietal topography with a prominent increase in power when participants closed their

eyes[24] (**Fig. 1e**). Auditory entrainment was dominant in fronto-central sensors with peaks in the ITC spectrum at the 3-Hz stimulus rate and its harmonic frequencies[19,23,25] (**Fig. 1f**).

2.2 α-oscillations and entrainment show slow fluctuations at similar time scales

Adapting an approach from Lakatos et al.[10] to human EEG, we traced the evolution of spontaneous α-power and auditory entrainment during the task (**Fig. 1a**). We used a sliding window approach to quantify how both of these measures change over time (see Materials and Methods).

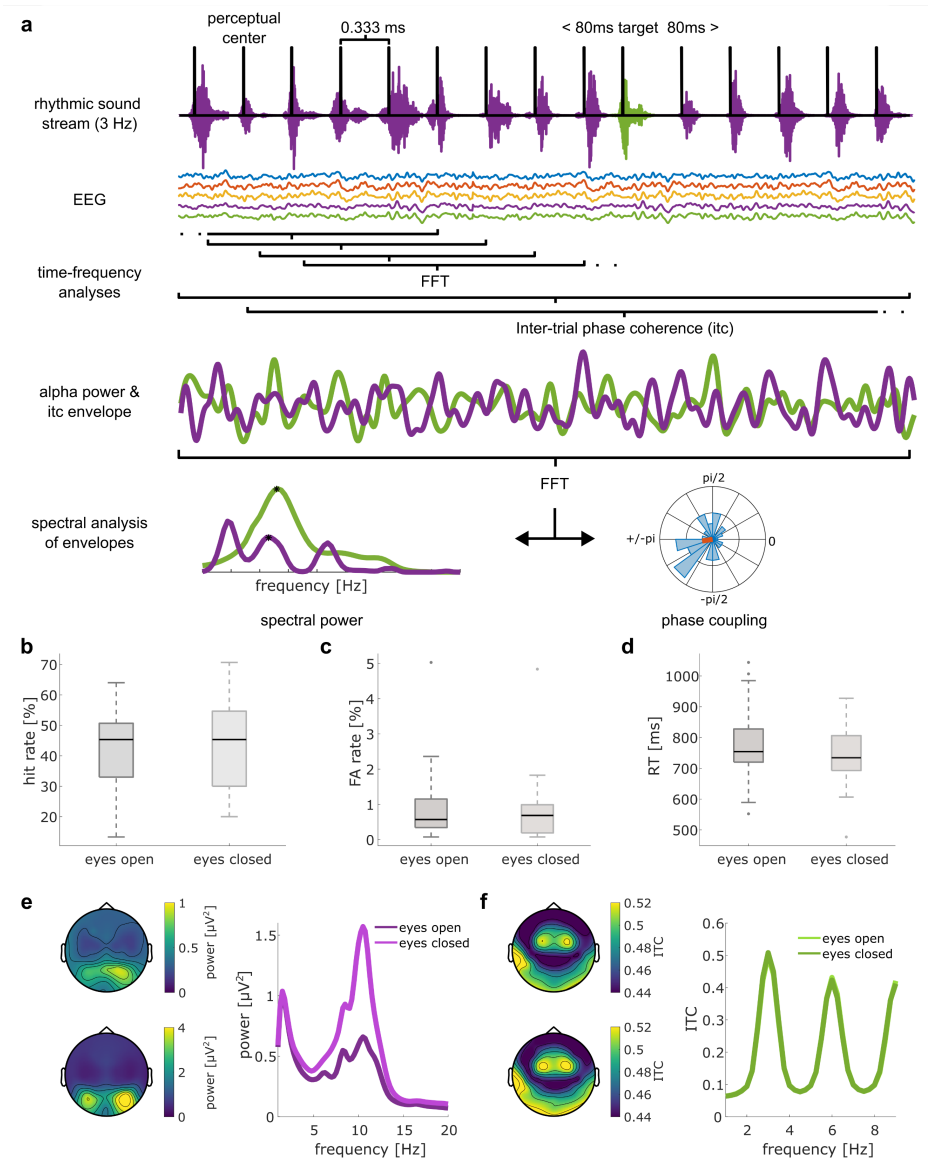

Figure 1: Experimental design & analysis. (a) Participants listened to continuous 5-min streams of rhythmic, monosyllabic French words presented at a rate of 3-Hz (top). Spectral analysis was performed on 2-s EEG segments centered on the perceptual center of each word. 15 adjacent segments (~5-sec window) were integrated in a sliding window approach to compute inter-trial coherence (ITC) over time. ITC at 3-Hz and power in the *α*-band (8 – 12-Hz) were extracted and treated as new time-series (4th row). The two time-series were submitted to another spectral analysis to assess slow, rhythmic fluctuations of *α*-oscillations and auditory entrainment. We identified prominent spectral peaks in both spectra and assessed their coupling and phasic relation (bottom row). **(b+c)** Proportion of hits and false alarms in eyes-open and eyes-open conditions. **(d)** Reaction times for hits. **(e)** Topography of *α*-power in eyes-open (top) and eyes-closed (bottom) conditions. Spectra on the right have been extracted from channel Pz and averaged across subjects. **(f)** Same as e, but for ITC (spectra are shown for channel Fz). Both *α*-power and ITC spectra and topographies are consistent with previous reports in the literature.

We found that both α-oscillations and neural entrainment exhibit slow, regular fluctuations **(Fig. 2 a)**. The dominant frequency in these fluctuations, which we revealed as 0.07-Hz (~14 sec, M_{alpha} = 0.0713-Hz ± *SD* = 0.0126, M_{ITC} = 0.0710-Hz ± *SD*

= 0.0116), was strikingly similar to that reported in non-human primates[10]. While the topographical distribution of entrainment fluctuations (**Fig. 2b**) resembled that of entrainment itself (**Fig. 1e**), this was not the case for α-oscillations. This result implies that the observed fluctuations in α-power (**Fig. 2b**) might be indeed linked to auditory processing, in contrast to the distribution of α-power that is generally dominated by the visual system (**Fig. 1d**).

α-oscillations and entrainment did not only fluctuate at similar time scales on the group level (**Fig. 2c**), but also within individuals: On average, the individual peak frequency for α-power fluctuations did not differ from that for neural entrainment (dependent samples t-test: t_{22} = 0.08, p = .93; $M_{|alpha - ITC|}$ = 0.015-Hz ± SD = 0.0112-Hz). Together, we found that α-oscillations and neural entrainment exhibit similar slow, regular fluctuations.

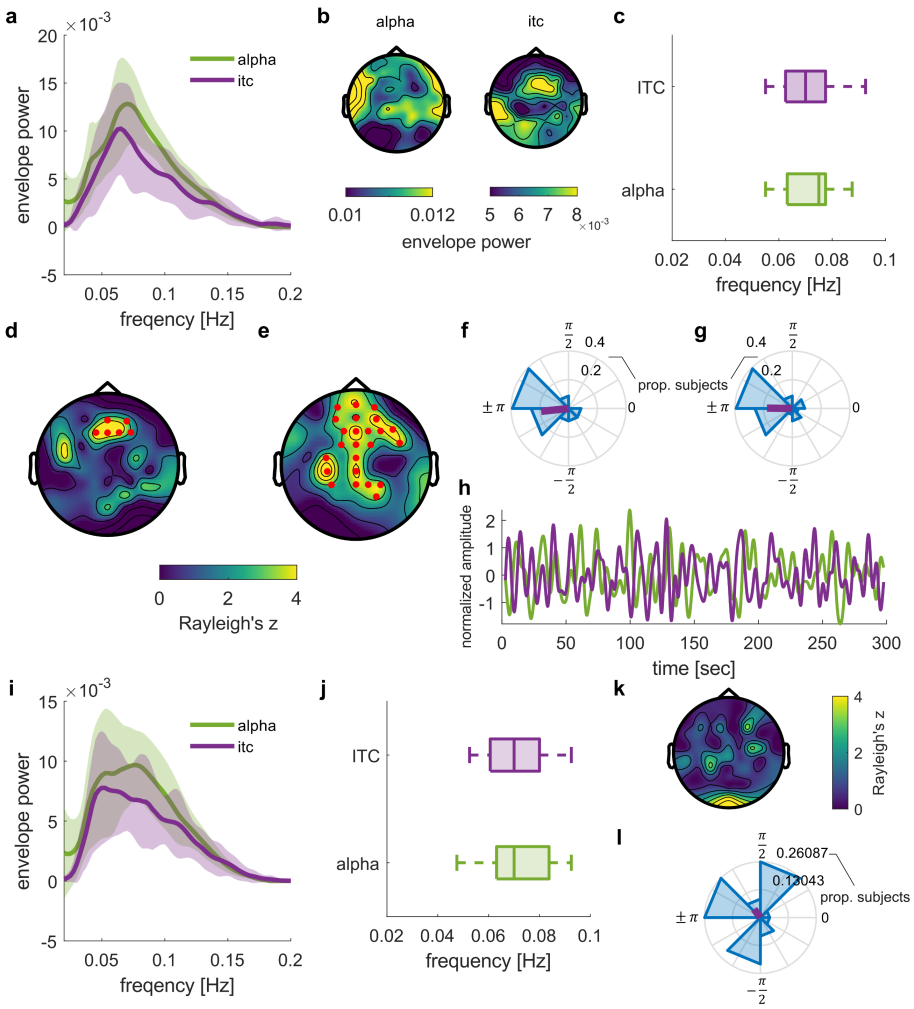

Figure 2: α-power and entrainment exhibit slow anti-phase fluctuations. (a-h) Results for eyes-open condition. **(a)** Envelope spectra of α-power and entrainment to rhythmic speech show a peak around 0.07-Hz (shown for electrode Fz). Shaded areas depict standard deviation. **(b)** Topography of the 0.07-Hz peak shown in a. **(c)** Distribution of individual peak frequencies from the spectra shown in a. **(d)** Coupling between α-power and entrainment fluctuations around 0.07-Hz when both were extracted from the same channel. Channels showing a significantly non-uniform distribution of phase differences are highlighted in red, topography indicates the underlying z-statistic. **(e)** Channels showing significant coupling between α-power fluctuations per channel with entrainment in the frontal cluster (d). **(f)** Distributions of phases in channels showing significant α-power vs. entrainment coupling (cluster shown in d). α-power and entrainment to speech are coupled in anti-phase. **(g)** α-power fluctuations in cluster shown in (e) are coupled in anti-phase. **(h)** Exemplary time-course of α-power and entrainment fluctuations at electrode Fz. **(i-l)** same as a, c, d, g, but for eyes-open condition.

2.3 Anti-phasic relation between α-power and entrainment fluctuations

We next assessed if the rhythmic fluctuations of α-oscillations and auditory entrainment are coupled. If the two reflected opposing processing modes, they should show an anti-phase relationship: When α-oscillations are strong, neural entrainment should be reduced, and vice versa. To this end, we computed the average phase difference between α-power and entrainment fluctuations for each EEG channel. Our analysis revealed a significant coupling (i.e., constant phase relation) between α-power and entrainment within a cluster of fronto-central channels (cluster-based Rayleigh's test; $p_{cluster}$ = .037, **Fig. 2d,** see **Supplementary Table 1** for an overview of channels within the cluster). A circular one-sample test yielded a significant deviation of the average phase difference from zero (p < .001). This average phase difference was close to anti-phase (M_{angle} = -3.04 rad), and with the 99% confidence interval for the sample mean including $\pm\pi$ (CI_{99} = 2.39, -2.18, **Fig. 2f**). **Supplementary Fig. S1a** provides an overview of the phase distribution in each EEG channel. Importantly, the anti-phasic relation of the two signals was evident on single subject level. 14 out of 23 participants show a significant coupling within the identified cluster. For 15 out of 23 participants, the average angle between α-power and entrainment fluctuations significantly differed from 0 (**Supplementary Fig. S1b**).

For results described above and shown in **Fig. 2d**, we contrasted α-power and entrainment from the same EEG channels. We next tested whether different channel combinations produce similar results. The topographical distribution of auditory entrainment in the EEG is well established and was reproduced in our results (**Figs. 2b,d**). However, α-oscillations are typically dominated by vision and their topographical pattern was more difficult to predict in our case. We therefore assessed, separately for each channel, whether α-power in this channel is coupled with neural entrainment in the frontal channel cluster shown in **Fig. 2d**. The analysis revealed a more distributed cluster of fronto-central and parietal channels in which α-power fluctuations are coupled to auditory entrainment (random permutation cluster Rayleigh-test: $p_{cluster}$ = .042, **Fig. 2e,** see **Supplementary Table 1** for an overview of channels within the cluster). Again, the difference between entrainment and α-power fluctuations was close to anti-phase within this cluster (M_{angle} = 3.13 rad, circular one-sample test against angle of zero: p < .01) with the 99% CI for the sample mean including $\pm\pi$ (CI_{99} = 2.16, -2.18, **Fig.**

2g). **Fig. 2h** depicts an exemplary time course of α-power and entrainment fluctuations. An example for each participant is shown in **Supplementary Fig. S1c**. In control analyses, we ruled out that the fluctuations observed in the α-band are driven by harmonic entrainment at frequencies in the α-band (**Supplementary Fig. S2**). Together, we found an anti-phase relation between the slow fluctuations in α-power and neural entrainment, as predicted from two opposing neural processes[9,12].

2.4 Anti-phase relation between entrainment and α-oscillations is state dependent

Interestingly, when participants were instructed to close their eyes during the task, the 0.07-Hz peaks became less pronounced, compared to the eyes-open condition (**Fig. 2i-l**). Further, we did not find evidence for a significant coupling ($p_{cluster}$ > .68, **Fig. 2k**) nor anti-phase relationship (circular one-sample test against angle of zero: p > .05, M_{angle} = 2.22 rad, **Fig. 2l**) between α-oscillations and entrainment when the eyes were closed.

2.5 Neural signatures of attentional mode differ between detected and missed target stimuli

Thus far we have reported a systematic coupling between slow fluctuations of α-power and auditory entrainment to rhythmic speech. Periods of stronger entrainment and lower α-power alternated with periods of weaker entrainment and higher α-power. If such periods are indeed indicative of different attentional processing modes, they should also be related to target detection. In other words, α-power (or entrainment) fluctuations prior to detected targets should be in opposite phase as compared to missed ones. **Fig. 3a**,**c** depicts the average time courses of α-power and entrainment fluctuations around hits and misses (bandpass filtered around the individual peak frequencies around 0.07-Hz). We computed the instantaneous phase of these time courses on an individual level. We observed a systematic clustering of phase differences between hits and misses for both α-power (Rayleigh test: p = 0.009, z = 4.61), and entrainment fluctuations (Rayleigh test: p = .049, z = 2.07) in the time period before a target occurred (-14-sec to -2.5-sec). Importantly, the mean angle of this difference significantly differed from 0 (circular one-sample test against angle of zero; alpha: p < 0.01, M_{angle} = -2.87, CI_{99} = 2.42, -1.88; entrainment: p < 0.05, M_{angle} = 2.44 rad, CI_{95} = 1.52, -2.93). The CI_{99} of this difference included $\pm\pi$ (**Fig. 3b,e)**, indicating an anti-phase relation**.** The observed

phase relations appeared to be relatively stable across the pre-stimulus time period (**Fig. 3c,f**). Together, we found that the hypothesized markers of attentional processing (entrainment vs. alpha) differ depending on whether an auditory target was detected or not, as expected from modes of external and internal attention, respectively.

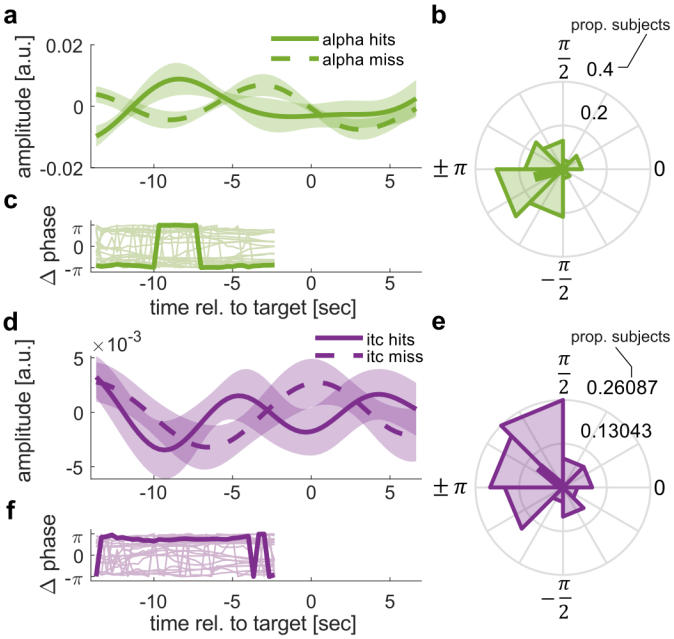

Figure 3: α-power and entrainment fluctuations differ between detected and missed targets. (**a**) Average, bandpass filtered *α*-power fluctuations around hits and misses extracted from significant cluster shown in Fig. 2e. Shaded areas indicate standard error of the mean. Note that data after -2.5 s can be affected by "smearing" of post—target data and thus cannot be interpreted in light of our hypothesis. (**b**) Polar histogram depicts the distribution of phase differences between *α*-power fluctuations prior to hits and misses, respectively (-14-sec to -2.5-sec). (**c**) Event-related phase difference of *α*-power fluctuations prior to hits and misses. Thin lines indicate single subject time-courses. Bold line depicts the circular average. Data is only shown for the time period used for statistical analysis. Note that -π = π, suggesting a stable phase opposition between envelopes preceding hits and misses, respectively. (**d-f**) Same as a-c, but for ITC.

3 Discussion

Variations in attentional performance are a prominent feature of sustained attention. In the current study, our marker of sustained attention to speech – neural entrainment – exhibited slow fluctuations with an inherently rhythmic component (**Fig. 2a**). Importantly, these fluctuations were opposite to those in α-oscillations (**Fig. 2f**), commonly assumed to reflect suppressed sensory input[13,14] and therefore indicative of an opposite mode of "internal" attention. In addition, neural signatures of attentional mode differed depending on whether a target was detected or not (**Fig. 3**). Our results therefore demonstrate that lapses in (external) attention occur rhythmically, even when presented with a stimulus that requires sustained attention for successful comprehension. Moreover, these fluctuations occurred at time scales (~14-sec) that are very similar to those observed in non-human primates[10]. Thus, we might have tapped into a general property of sustained attention that is conserved across species.

An anti-phase relation is consistent with the general roles commonly ascribed to α-oscillations and auditory entrainment. Pronounced auditory entrainment indicates an attentional focus on external stimulus processing and has been found to be beneficial to auditory perception and speech comprehension[20,21,26–28]. Conversely, α-oscillations have been linked to inhibition of sensory

processing[13,14,29,30] and attentional (de-)selection[15–18,31]. In the context of sustained attention, previous work has linked attentional lapses to pre-stimulus levels of α-oscillations[11,32], and periods of mind wandering[33].

It remains unclear why the observed effects vanish when participants close their eyes. The dominantly frontal coupling between α-power and auditory entrainment may indicate that these originate from temporal auditory, and possibly parietal attention related networks. However, closing the eyes is known to cause a substantial increase of α-oscillations pre-dominantly in visual areas[24]. It is possible that these enhanced visual α-oscillations overshadow their auditory and parietal counterparts when eyes are closed, such that their coupling to auditory entrainment cannot be traced anymore. Alternatively, eye-closure may cause a fundamental change in the brain's processing mode. Blocking visual input may allow to allocate more cognitive resources to auditory processing, such that rhythmic switching may occur at fundamentally different frequencies or is not necessary at all.

Our results pose important questions about the putative mechanisms driving the remarkable rhythmicity of attentional fluctuations observed in our data and their function. An obvious concern might be that the rhythmicity is inherently driven by the regularity of the stimulus material. However, this seems unlikely. Low frequency effects caused by the rhythmicity of the stimulus material should follow the principles of synchronization theory, which would predict such effects to occur at precise, predictable subharmonic frequencies[34], while the fluctuations we observe in our data vary across participants. It therefore seems likely that the observed fluctuations are intrinsically driven. Indeed, at the level of short sub-second time-scales it has been repeatedly suggested that perceptual and attentional sampling are inherently rhythmic, fluctuating in the range of theta and α-oscillations[35–42]. It may thus be plausible that rhythmicity in attention can also exists on other time scales. We speculate that continuous fluctuations between internal and external attentional modes might implement a protective mechanism to prevent depletion of attentional resources. Indeed, evidence from non-invasive brain stimulation suggests that enhancing endogenous α-oscillations with electrical stimulation has a stabilizing effect on sustained attention[43].

There is an intriguing similarity of timescales between the intrinsic attentional rhythms in our data and the known coupling of α-oscillations to slow fluctuations of extra-cerebral sources in the body

such as the respiratory system[44], the heart, and the gut[45]. In particular, the gastric network seems to generate rhythmic activity at similar time scales as those we found here (~0.05 Hz)[45,46], and modulate the amplitude of spontaneous alpha-oscillations[47]. Future research should address if these similarities in time scales are a mere co-incidence or if they may be involved in driving slow attentional fluctuations.

A systematic rhythmicity of attentional fluctuations has important implications for both basic and applied research. Considering rhythmicity may make attentional lapses more predictable and offer a potential target for interventional approaches. For example, transcranial alternating current stimulation (tACS) can be used to modulate brain oscillations[48–50] and has been previously shown to stabilize sustained attention when applied in a continuous manner[43]. Considering fluctuations of attentional modes may allow to apply tACS in a state-dependent manner, e.g., to induce shifts in the attentional state by applying stimulation either in the α-frequency range or in synchrony with the external stimulus. Such targeted intervention may offer novel opportunities to improve or steer sustained attention performance in critical systems or in neurological or psychiatric patients suffering from deficits in sustained attention[7,8]. Before moving to such practical applications, additional research should investigate to what extent the current findings generalize across sensory systems. Compared to other sensory domains with a more static input, audition is special in that information is inherently transient and may thus benefit more from processing principles that take its temporal structure into account[51]. It thus remains to be determined if similar rhythmicity exists when sustained attention is deployed to visual, somatosensory, or cross-modal tasks.

4 Materials & Methods

4.1 Participants

Twenty-three healthy volunteers (age 22.4 years ± 1.6 years, 15 females) participated in the study. They gave written informed consent prior to joining the experiment and were paid for their participation. The study was approved by the ethics board CPP (Comité de Protection des Personnes) Ouest II Angers (proposal no: CPP 21.01.22.71950 / 2021-A00131-40).

4.2 Experimental Design

Over the course of six 5-min blocks, participants were instructed to listen to continuous streams of rhythmic, one-syllable French words and to indicate if they detected deviations from the rhythm via a button press on a standard computer keyboard. In the beginning of each block, they were instructed to either keep their eyes-open and fixated on a white cross at the center of a computer screen, or to keep them closed. Half the blocks were assigned to the eyes-open, and eyes-closed conditions respectively. The order of blocks was randomized to avoid time-on task effects. Participants were familiarized with the task prior to the main experiment. They were shown examples of continuous rhythmic speech trains as well as streams containing violations of the rhythm. Subsequently, they performed a 1-min practice run of the task.

4.3 Apparatus and stimuli

Original recordings consisted of a set of 474 monosyllabic French words, spoken to a metronome at a rate of 2-Hz by a male, native French speaker. This approach aligned perceptual centers (p-centers)[52] of the words to the metronome beat and resulted in perceptually very rhythmic speech (see Zoefel et al[53] for a detailed description of stimuli and task). Stimuli were then time-compressed to 3-Hz using the pitch-synchronous overlap and add (PSOLA) algorithm implemented in the Praat software package, and the metronome beat was made inaudible to participants. Intelligibility of the speech recordings was degraded by applying 16-channel noise-vocoding[54]. 16-channel noise-vocoded speech is not as easy to understand as clear speech, but is still clearly intelligible[55]. Individual, noise-vocoded words were then concatenated into a continuous sound stream of 5-min length. The order of words was randomized with a constraint such that every word from the stimulus set had

occurred before it can be repeated. 150 target words that deviated from the 3-Hz stimulus rate by 80-ms (50% presented early, 50% late) were embedded into the rhythmic speech stream. Participants were asked to indicate violations of the rhythm by pressing the space bar on a standard computer keyboard.

Audio signals were generated in MATLAB 2019a and streamed to a Fireface UCX (RME Audio, Heimhausen, Germany) soundcard. The audio stream was presented to participants using an In-Ear headphone system (Etymotic Research ER-2, Etymotic Research Inc., USA). Experimental instructions were given via a computer screen in the experimental room controlled using Psychtoolbox 3 for MATLAB. During blocks that required participants to keep their eyes-open, a white fixation cross was presented at the center of the screen.

4.4 EEG

Electroencephalogram was recorded from 64 active electrodes according to the extended international 10-10 system, using a BioSemi Active 2 amplifier (BioSemi, Amsterdam, Netherlands). EEG signals were recorded at a rate of 2048-Hz and digitally stored on a hard drive using ActiView v9.02 Software (BioSemi, Amsterdam, Netherlands). Electrodes were mounted in an elastic cap and connected to participants' scalps via a conductive gel (Signa Gel, Parker Laboratories Inc., Fairfield, NJ, USA). Signal offsets of the system were kept below 50-µV.

4.5 Data analysis

4.5.1 EEG processing

EEG analyses were performed in MATLAB 2019b using the fieldtrip toolbox[56]. Data was re-referenced to common average, resampled to 256-Hz and filtered between 1-Hz and 40-Hz using a two-pass, 4-th order, zero-phase Butterworth filter. An independent component analysis was performed to project out artifacts related to eye-blinks, movements, heart-beat or muscular activity.

Signals were then epoched into consecutive, overlapping segments centered around the p-center (the part of the word that was centered on the metronome beat) of each word (± 1-sec). Each segment was 2-sec long and therefore comprised seven p-centers (**Fig. 1**). The use of segments allowed us to extract time-resolved measures of α-oscillations and neural entrainment.

A Fast-Fourier Transform (FFT, Hanning window, 2-sec zero padding) was applied on each of the segments. The resulting complex Fourier coefficients were used to extract power in the α-band (8-12-Hz) as well as inter-trial coherence (ITC) at the stimulus rate (3-Hz). In line with previous work[10,23], we used ITC to quantify neural entrainment to speech, i.e. neural activity aligned to the 3-Hz rhythm. ITC quantifies phase consistency across trials (here: segments). ITC was computed in sliding windows comprising 15 segments (step size: 1 segment). Note that, due to the overlap between successive segments, this window is 5 seconds long. We used the following equation to compute ITC in each time window:

$$ITC(f) = \left| \frac{1}{N} \sum_{n=1}^{N} e^{-i(\varphi(f,n))} \right|$$

where $\varphi(f,n)$ is the phase in segment n at frequency f. N corresponds to the number of segments in the window. *f* was therefore set to 3-Hz and *N* was set to 15. Within the same windows we averaged power spectra across segments to ensure consistent temporal smoothing in both measures.

This approach yielded neural measures as a function of time: One α-power, and one ITC value per time window. We then used these time-resolved measures to extract their fluctuations over time. ITC and α-power time-series were first z-transformed to ensure comparable amplitudes. They were then divided into 100-sec segments with 90% overlap. This resulted in a total of 60 segments across the 3 blocks per condition (eyes-open vs. eyes-closed). Finally, the segments were submitted to another FFT (hanning window, 400-sec zero padding). Length of the segments and padding were chosen to ensure sufficient spectral resolution below 0.2-Hz as well as a sufficient number of phase estimates to quantify if there is systematic coupling between signals. Importantly, a 90% overlap results in an effective step size of 10-sec in our case. Given our aim to analyze fluctuations at rates below 0.2-Hz (i.e., slower than 5-sec), a 10-sec step size is sufficient to include 0.5 – 2 new cycles of such slow fluctuations. To rule out that these step and window size parameters affect our results we repeated the analysis using shorter segments (50-sec with 90% overlap, i.e., 5-sec time-steps). Overall, we obtained very similar results compared to the main analysis.

The obtained low-frequency spectra were corrected for arrhythmic ("1/f") activity using the fooof algorithm[57] as shipped with the fieldtrip toolbox. This step helped improve the identification of peaks

in the spectrum, and revealed prominent spectral peaks in the α-power and ITC time-courses. On the group level, these peaks were close to 0.07-Hz for both α-power and ITC (**Fig. 2a**). We next identified individual peak frequencies of α-power and ITC fluctuations. To do so, we selected the peak frequency that was closest to 0.07-Hz rhythm for individual participants (in a range of 0.04-Hz to 0.1-Hz). In accordance with Lakatos et al 2016[10], we then used the individual peak frequency of the α-power envelope to extract the phase of α-power and ITC fluctuations in each 100-sec segment, and computed their phase difference. Phase differences were subsequently averaged across segments, separately for each subject and condition. To rule out that coupling between α-power and entrainment are driven by entrainment effects at harmonic frequencies in the α-band (9-Hz & 12-Hz), we repeated the analysis for the coupling between fluctuations in 3-Hz ITC and ITC at 9-Hz and 12-Hz. The results of this analysis are presented in **Supplementary Fig. S2**.

To investigate how α-power and ITC fluctuations relate to the detection of target stimuli, we filtered the corresponding envelopes around the individual α-power envelope peak frequency (± 0.02-Hz), identified in the previous step, using a causal, 6-th order, one-pass Butterworth filter. The causal filter was chosen to avoid contamination of pre-stimulus activity with stimulus related changes in brain activity. We epoched the signals from -14-sec to +7-sec around target stimuli. Using a Hilbert transform, we then extracted instantaneous phase angles over time and averaged them across trials, separately for ITC and α-power fluctuations around hits and misses, respectively. Subsequently, for both alpha power and ITC fluctuations, we computed the time-resolved phase difference between hits and misses in the interval before target onset (-14-sec to -2.5-sec). The interval ends 2.5-sec prior to the onset of target stimuli to avoid including target-evoked brain responses, which can smear into the interval due to the symmetric 5-sec window (described above) that was used to compute ITC values and to smoothen α-power trajectories. The analysis was restricted to channels from significant clusters revealed in previous analyses (**Fig. 2d,e**; **Supplementary table 1**).

4.5.2 Behavioral analysis

Average reaction times, as well as the proportion of hits, misses and false alarms were computed for each condition (eyes-open vs. eyes-closed). A target word was considered a hit, if a button was pressed within 2-sec after stimulus onset, otherwise it was considered a miss. Button presses

occurring outside these response intervals were considered false-alarms. Hit rates were computed by dividing the number of responses to targets by the number of targets (75 per condition), false alarm rates were computed by dividing responses outside of response intervals by the number of standards (2475 per condition).

4.5.3 Statistical analyses

Statistical analyses were performed in Matlab 2019b using the circular statistics toolbox[58] in combination with functions in fieldtrip for massive-multivariate permutation statistics[56,59].

To assess whether there is a significant coupling between α-power and entrainment fluctuations (i.e. envelopes), we tested if their phase difference shows a systematic clustering across subjects (i.e. differs from a uniform distribution). To this end, we first computed the average phase difference between α-power and ITC envelopes (separately for each channel and subject as detailed in section 4.5.1) and subjected them to Rayleigh's test for uniformity of circular data. This test yields a z-statistic (one per channel) which is high if these phase differences are non-uniformly distributed across subjects. We next tested whether there is a cluster of channels with such a non-uniform distribution. We randomly shuffled α-power and ITC segments within subjects 10,000 times, re-computed the average phase difference between the two signals per-subject and the corresponding Rayleigh's statistic on the group level, yielding 10,000 z-statistics per channel. Finally, we compared actual data with shuffled ones to obtain group-level p-values for channel clusters, using Monte Carlo estimates[59].

A significantly non-uniform distribution of phase differences between α-power and ITC envelopes indicates that the two are coupled, but does not tell us anything about their phase relation. If this phase relation is close to 0, this would speak against alpha oscillations and entrainment reflecting opposing processing modes. We therefore tested the phase relation between the two deviates from zero within an identified channel cluster, using the circ_mtest function of the circular statistics toolbox. The function tests if a given angle lies within a given confidence interval (e.g., 95% or 99%) of a phase distribution. To estimate a p-value, we performed the test against different significance levels ($\alpha < .05$ and $\alpha < .01$) and report the lowest significant alpha level for each comparison (i.e., $p < .05$ or $p < .01$), along with the circular average of the underlying phase distribution and the 95% or 99% circular confidence intervals (CI). Although a non-zero phase relation between alpha oscillations

and entrainment might already indicate different modes of processing, our hypothesis was more explicit in that it assumes an opposing (i.e. anti-phase) relation between the two. To evaluate this hypothesis, we also report if the CIs cover ±π.

Comparisons of peak frequencies, hit and false alarm rates and reaction times were performed in Matlab 2019b using dependent-samples t-tests.

5 Data & code availability

The underlying data and code are available from the corresponding author upon reasonable request.

6 Acknowledgements:

This research was supported by a grant from the Agence Nationale de la Recherche (grant number ANR-21-CE37-0002) awarded to Benedikt Zoefel.

7 Autor contributions

Florian H. Kasten: Conceptualization, Methodology, Software, Data Collection, Formal Analysis, Investigation, Writing – Original Draft, Visualization, **Quentin Busson:** Data Collection, Software, Writing – Review & Editing, **Benedikt Zoefel:** Conceptualization, Writing – Review & Editing, Supervision, Funding Acquisition

8 Conflict of interest:

FHK, QB, and BZ declare no competing interests.

9 References

1. Esterman, M., Noonan, S.K., Rosenberg, M., and DeGutis, J. (2013). In the Zone or Zoning Out? Tracking Behavioral and Neural Fluctuations During Sustained Attention. Cerebral Cortex *23*, 2712–2723. 10.1093/cercor/bhs261.

2. Weissman, D.H., Roberts, K.C., Visscher, K.M., and Woldorff, M.G. (2006). The neural bases of momentary lapses in attention. Nat Neurosci *9*, 971–978. 10.1038/nn1727.

3. deBettencourt, M.T., Norman, K.A., and Turk-Browne, N.B. (2018). Forgetting from lapses of sustained attention. Psychon Bull Rev *25*, 605–611. 10.3758/s13423-017-1309-5.

4. Edkins, G.D., and Pollock, C.M. (1997). The influence of sustained attention on Railway accidents. Accident Analysis & Prevention *29*, 533–539. 10.1016/S0001-4575(97)00033-X.

5. Taylor-Phillips, S., Elze, M.C., Krupinski, E.A., Dennick, K., Gale, A.G., Clarke, A., and Mello-Thoms, C. (2015). Retrospective Review of the Drop in Observer Detection Performance Over Time in Lesion-enriched Experimental Studies. J Digit Imaging *28*, 32–40. 10.1007/s10278-014-9717-9.

6. Schwebel, D.C., Lindsay, S., and Simpson, J. (2007). Brief Report: A Brief Intervention to Improve Lifeguard Surveillance at a Public Swimming Pool. Journal of Pediatric Psychology *32*, 862–868. 10.1093/jpepsy/jsm019.

7. Gmehlin, D., Fuermaier, A.B.M., Walther, S., Tucha, L., Koerts, J., Lange, K.W., Tucha, O., Weisbrod, M., and Aschenbrenner, S. (2016). Attentional Lapses of Adults with Attention Deficit Hyperactivity Disorder in Tasks of Sustained Attention. Archives of Clinical Neuropsychology *31*, 343–357. 10.1093/arclin/acw016.

8. Greer, J., Riby, D.M., Hamiliton, C., and Riby, L.M. (2013). Attentional lapse and inhibition control in adults with Williams Syndrome. Research in Developmental Disabilities *34*, 4170–4177. 10.1016/j.ridd.2013.08.041.

9. Chun, M.M., Golomb, J.D., and Turk-Browne, N.B. (2011). A Taxonomy of External and Internal Attention. Annual Review of Psychology *62*, 73–101. 10.1146/annurev.psych.093008.100427.

10. Lakatos, P., Barczak, A., Neymotin, S.A., McGinnis, T., Ross, D., Javitt, D.C., and O'Connell, M.N. (2016). Global dynamics of selective attention and its lapses in primary auditory cortex. Nat Neurosci *19*, 1707–1717. 10.1038/nn.4386.

11. O'Connell, R.G., Dockree, P.M., Robertson, I.H., Bellgrove, M.A., Foxe, J.J., and Kelly, S.P. (2009). Uncovering the Neural Signature of Lapsing Attention: Electrophysiological Signals Predict Errors up to 20 s before They Occur. J. Neurosci. *29*, 8604–8611. 10.1523/JNEUROSCI.5967-08.2009.

12. Clayton, M.S., Yeung, N., and Kadosh, R.C. (2015). The roles of cortical oscillations in sustained attention. Trends in Cognitive Sciences *19*, 188–195. 10.1016/j.tics.2015.02.004.

13. Jensen, O., and Mazaheri, A. (2010). Shaping Functional Architecture by Oscillatory Alpha Activity: Gating by Inhibition. Front. Hum. Neurosci. *4*. 10.3389/fnhum.2010.00186.

14. Klimesch, W., Sauseng, P., and Hanslmayr, S. (2007). EEG alpha oscillations: The inhibition–timing hypothesis. Brain Research Reviews *53*, 63–88. 10.1016/j.brainresrev.2006.06.003.

15. Kasten, F.H., Wendeln, T., Stecher, H.I., and Herrmann, C.S. (2020). Hemisphere-specific, differential effects of lateralized, occipital–parietal α- versus γ-tACS on endogenous but not exogenous visual-spatial attention. Sci Rep *10*, 12270. 10.1038/s41598-020-68992-2.

16. Haegens, S., Handel, B.F., and Jensen, O. (2011). Top-Down Controlled Alpha Band Activity in Somatosensory Areas Determines Behavioral Performance in a Discrimination Task. Journal of Neuroscience *31*, 5197–5204. 10.1523/JNEUROSCI.5199-10.2011.

17. Okazaki, Y.O., De Weerd, P., Haegens, S., and Jensen, O. (2014). Hemispheric lateralization of posterior alpha reduces distracter interference during face matching. Brain Research *1590*, 56–64. 10.1016/j.brainres.2014.09.058.

18. Wöstmann, M., Herrmann, B., Maess, B., and Obleser, J. (2016). Spatiotemporal dynamics of auditory attention synchronize with speech. Proc Natl Acad Sci USA *113*, 3873–3878. 10.1073/pnas.1523357113.

19. Cabral-Calderin, Y., and Henry, M.J. (2022). Reliability of Neural Entrainment in the Human Auditory System. J. Neurosci. *42*, 894–908. 10.1523/JNEUROSCI.0514-21.2021.

20. Henry, M.J., and Obleser, J. (2012). Frequency modulation entrains slow neural oscillations and optimizes human listening behavior. Proceedings of the National Academy of Sciences *109*, 20095–20100. 10.1073/pnas.1213390109.

21. Riecke, L., Formisano, E., Sorger, B., Başkent, D., and Gaudrain, E. (2018). Neural Entrainment to Speech Modulates Speech Intelligibility. Current Biology *28*, 161-169.e5. 10.1016/j.cub.2017.11.033.

22. Pöppel, E. (1997). A hierarchical model of temporal perception. Trends in Cognitive Sciences *1*, 56–61. 10.1016/S1364-6613(97)01008-5.

23. van Bree, S., Sohoglu, E., Davis, M.H., and Zoefel, B. (2021). Sustained neural rhythms reveal endogenous oscillations supporting speech perception. PLOS Biology *19*, e3001142. 10.1371/journal.pbio.3001142.

24. Barry, R.J., Clarke, A.R., Johnstone, S.J., Magee, C.A., and Rushby, J.A. (2007). EEG differences between eyes-closed and eyes-open resting conditions. Clinical Neurophysiology *118*, 2765–2773. 10.1016/j.clinph.2007.07.028.

25. Pérez, A., Davis, M.H., Ince, R.A.A., Zhang, H., Fu, Z., Lamarca, M., Lambon Ralph, M.A., and Monahan, P.J. (2022). Timing of brain entrainment to the speech envelope during speaking, listening and self-listening. Cognition *224*, 105051. 10.1016/j.cognition.2022.105051.

26. Giraud, A.-L., and Poeppel, D. (2012). Cortical oscillations and speech processing: emerging computational principles and operations. Nat Neurosci *15*, 511–517. 10.1038/nn.3063.

27. Henry, M.J., Herrmann, B., and Obleser, J. (2014). Entrained neural oscillations in multiple frequency bands comodulate behavior. Proceedings of the National Academy of Sciences *111*, 14935–14940. 10.1073/pnas.1408741111.

28. Zoefel, B., Archer-Boyd, A., and Davis, M.H. (2018). Phase Entrainment of Brain Oscillations Causally Modulates Neural Responses to Intelligible Speech. Current Biology *28*, 401-408.e5. 10.1016/j.cub.2017.11.071.

29. Hanslmayr, S., Aslan, A., Staudigl, T., Klimesch, W., Herrmann, C.S., and Bäuml, K.-H. (2007). Prestimulus oscillations predict visual perception performance between and within subjects. NeuroImage *37*, 1465–1473. 10.1016/j.neuroimage.2007.07.011.

30. van Dijk, H., Schoffelen, J.-M., Oostenveld, R., and Jensen, O. (2008). Prestimulus Oscillatory Activity in the Alpha Band Predicts Visual Discrimination Ability. Journal of Neuroscience *28*, 1816–1823. 10.1523/JNEUROSCI.1853-07.2008.

31. Händel, B.F., Haarmeier, T., and Jensen, O. (2011). Alpha Oscillations Correlate with the Successful Inhibition of Unattended Stimuli. Journal of Cognitive Neuroscience *23*, 2494–2502. 10.1162/jocn.2010.21557.

32. Boudewyn, M.A., and Carter, C.S. (2018). I must have missed that: Alpha-band oscillations track attention to spoken language. Neuropsychologia *117*, 148–155. 10.1016/j.neuropsychologia.2018.05.024.

33. Compton, R.J., Gearinger, D., and Wild, H. (2019). The wandering mind oscillates: EEG alpha power is enhanced during moments of mind-wandering. Cogn Affect Behav Neurosci *19*, 1184–1191. 10.3758/s13415-019-00745-9.

34. Pikovsky, A., Rosenblum, M., and Kurths, J. (2003). Synchronization: A universal concept in nonlinear sciences (Cambridge University Press).

35. Helfrich, R.F., Fiebelkorn, I.C., Szczepanski, S.M., Lin, J.J., Parvizi, J., Knight, R.T., and Kastner, S. (2018). Neural Mechanisms of Sustained Attention Are Rhythmic. Neuron *99*, 854-865.e5. 10.1016/j.neuron.2018.07.032.

36. Busch, N.A., Dubois, J., and VanRullen, R. (2009). The Phase of Ongoing EEG Oscillations Predicts Visual Perception. Journal of Neuroscience *29*, 7869–7876. 10.1523/JNEUROSCI.0113-09.2009.

37. VanRullen, R., and Koch, C. (2003). Is perception discrete or continuous? Trends in Cognitive Sciences *7*, 207–213. 10.1016/S1364-6613(03)00095-0.

38. Kasten, F.H., and Herrmann, C.S. (2022). Discrete sampling in perception via neuronal oscillations—Evidence from rhythmic, non-invasive brain stimulation. European Journal of Neuroscience *55*, 3402–3417. 10.1111/ejn.15006.

39. Fiebelkorn, I.C., and Kastner, S. (2019). A Rhythmic Theory of Attention. Trends in Cognitive Sciences *23*, 87–101. 10.1016/j.tics.2018.11.009.

40. Fiebelkorn, I.C., Saalmann, Y.B., and Kastner, S. (2013). Rhythmic Sampling within and between Objects despite Sustained Attention at a Cued Location. Current Biology *23*, 2553–2558. 10.1016/j.cub.2013.10.063.

41. Re, D., Inbar, M., Richter, C.G., and Landau, A.N. (2019). Feature-Based Attention Samples Stimuli Rhythmically. Current Biology *29*, 693-699.e4. 10.1016/j.cub.2019.01.010.

42. Landau, A.N., and Fries, P. (2012). Attention Samples Stimuli Rhythmically. Current Biology *22*, 1000–1004. 10.1016/j.cub.2012.03.054.

43. Clayton, M.S., Yeung, N., and Cohen Kadosh, R. (2019). Electrical stimulation of alpha oscillations stabilizes performance on visual attention tasks. Journal of Experimental Psychology: General *148*, 203–220. 10.1037/xge0000502.

44. Kluger, D.S., and Gross, J. (2021). Respiration modulates oscillatory neural network activity at rest. PLOS Biology *19*, e3001457. 10.1371/journal.pbio.3001457.

45. Azzalini, D., Rebollo, I., and Tallon-Baudry, C. (2019). Visceral Signals Shape Brain Dynamics and Cognition. Trends in Cognitive Sciences *23*, 488–509. 10.1016/j.tics.2019.03.007.

46. Wolpert, N., Rebollo, I., and Tallon-Baudry, C. (2020). Electrogastrography for psychophysiological research: Practical considerations, analysis pipeline, and normative data in a large sample. Psychophysiology *57*, e13599. 10.1111/psyp.13599.

47. Richter, C.G., Babo-Rebelo, M., Schwartz, D., and Tallon-Baudry, C. (2017). Phase-amplitude coupling at the organism level: The amplitude of spontaneous alpha rhythm fluctuations varies with the phase of the infra-slow gastric basal rhythm. NeuroImage *146*, 951–958. 10.1016/j.neuroimage.2016.08.043.

48. Kasten, F.H., Dowsett, J., and Herrmann, C.S. (2016). Sustained Aftereffect of α-tACS Lasts Up to 70 min after Stimulation. Front. Hum. Neurosci. *10*. 10.3389/fnhum.2016.00245.

49. Kasten, F.H., Duecker, K., Maack, M.C., Meiser, A., and Herrmann, C.S. (2019). Integrating electric field modeling and neuroimaging to explain inter-individual variability of tACS effects. Nat Commun *10*, 5427. 10.1038/s41467-019-13417-6.

50. Veniero, D., Vossen, A., Gross, J., and Thut, G. (2015). Lasting EEG/MEG Aftereffects of Rhythmic Transcranial Brain Stimulation: Level of Control Over Oscillatory Network Activity. Front. Cell. Neurosci. *9*. 10.3389/fncel.2015.00477.

51. VanRullen, R., Zoefel, B., and Ilhan, B. (2014). On the cyclic nature of perception in vision versus audition. Philosophical Transactions of the Royal Society B: Biological Sciences *369*, 20130214. 10.1098/rstb.2013.0214.

52. Scott, S.K. (1998). The point of P-centres. Psychological Research Psychologische Forschung *61*, 4–11. 10.1007/PL00008162.

53. Zoefel, B., Gilbert, R.A., and Davis, M.H. (2023). Intelligibility improves perception of timing changes in speech. PLOS ONE *18*, e0279024. 10.1371/journal.pone.0279024.

54. Shannon, R.V., Zeng, F.G., Kamath, V., Wygonski, J., and Ekelid, M. (1995). Speech recognition with primarily temporal cues. Science *270*, 303–304. 10.1126/science.270.5234.303.

55. Davis, M.H., Johnsrude, I.S., Hervais-Adelman, A., Taylor, K., and McGettigan, C. (2005). Lexical information drives perceptual learning of distorted speech: evidence from the comprehension of noise-vocoded sentences. J Exp Psychol Gen *134*, 222–241. 10.1037/0096-3445.134.2.222.

56. Oostenveld, R., Fries, P., Maris, E., and Schoffelen, J.-M. (2011). FieldTrip: Open Source Software for Advanced Analysis of MEG, EEG, and Invasive Electrophysiological Data. Computational Intelligence and Neuroscience *2011*, 1–9. 10.1155/2011/156869.

57. Donoghue, T., Haller, M., Peterson, E.J., Varma, P., Sebastian, P., Gao, R., Noto, T., Lara, A.H., Wallis, J.D., Knight, R.T., et al. (2020). Parameterizing neural power spectra into periodic and aperiodic components. Nature Neuroscience *23*, 1655–1665. 10.1038/s41593-020-00744-x.

58. Berens, P. (2009). CircStat: A MATLAB Toolbox for Circular Statistics. Journal of Statistical Software *31*, 1–21. 10.18637/jss.v031.i10.

59. Maris, E., and Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. Journal of Neuroscience Methods *164*, 177–190. 10.1016/j.jneumeth.2007.03.024.