# Opposing neural processing modes alternate rhythmically during sustained auditory attention

Florian Kasten, Quentin Busson, Benedikt Zoefel

# Opposing neural processing modes alternate rhythmically during sustained auditory attention

Florian H. Kasten[1,2,3*], Quentin Busson[3] and Benedikt Zoefel[2,3,*]

[1]Department for Cognitive, Affective, Behavioral Neuroscience with Focus Neurostimulation, Institute of Psychology, University of Trier, Trier, Germany

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

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

*Corresponding authors:

Benedikt Zoefel, benedikt.zoefel@cnrs.fr (phone: +33 5 62 74 61 62)

CNRS CERCO UMR 5549
Pavillon Baudot CHU Purpan
BP 25202
31052 Toulouse Cedex,
France

Florian H. Kasten, kasten.florian@cnrs.fr (phone: +49 651 201 3250)

Universität Trier
Universitätsring 15
54296 Trier
Germany

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

(Abstract: 191 words, Manuscript: 3929 words, Methods: 2290 words, Number of Figures: 5)

**Abstract**

During 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.

In the current study, we extract presumed neural signatures of these attentional modes in human electroencephalography (EEG): neural entrainment and α-oscillations (~10-Hz), linked to the processing and suppression of sensory information, respectively. We test whether they exhibit structured fluctuations over time, while listeners attend to an ecologically relevant stimulus, like speech, and complete a task that requires full and continuous attention.

Results show an antagonistic relation between neural entrainment to speech and spontaneous α-oscillations in two distinct brain networks – one specialized in the processing of external information, the other reminiscent of the dorsal attention network. These opposing neural modes undergo slow, periodic fluctuations around ~0.07 Hz and are related to the 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.

## 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, 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 these 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. When listeners attend to rhythmic sequences, neural activity synchronizes to the stimulus rhythm, an effect often termed neural entrainment[19–21]. As a potential counterpart to α-oscillations, neural entrainment is therefore a prime example for a marker of active stimulus processing[20,22,23]. However, neural entrainment and α-oscillations are often studied in isolation, and the few efforts to relate them[24–26] have used trial-based paradigms that, due to their frequent interruptions of stimulus presentation, cannot capture fluctuations in sustained attention.

One of the rare studies that did consider both of these neural markers during sustained attention measured them in the auditory cortex of macaque monkeys[10]. This work 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[19,25,26]. 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. Human speech requires integration of information over time to be optimally perceived[27]. It therefore needs full and continuous attention, and regular attentional lapses seem particularly harmful for speech processing. Finally, the network structure governing the hypothesized modes in auditory sustained attention remained unaddressed thus far.

In the current study, we recorded electroencephalographic (EEG) data in humans and tested for regular fluctuations in attentional modes while participants paid sustained attention to rhythmic speech sounds. We hypothesized that, similar to the aforementioned findings in non-human primates[10], neural entrainment to speech and spontaneous α-oscillations show rhythmic fluctuations at ultra-slow frequencies close to ~0.06 Hz (0.02-Hz - 0.2-Hz) and that these fluctuations show an antagonistic relationship (i.e. are coupled in anti-phase). Given its strong environmental rhythms[28], we here chose to focus on the auditory modality. However, our findings may pave the way for investigations into rhythmic attentional fluctuations across sensory modalities.

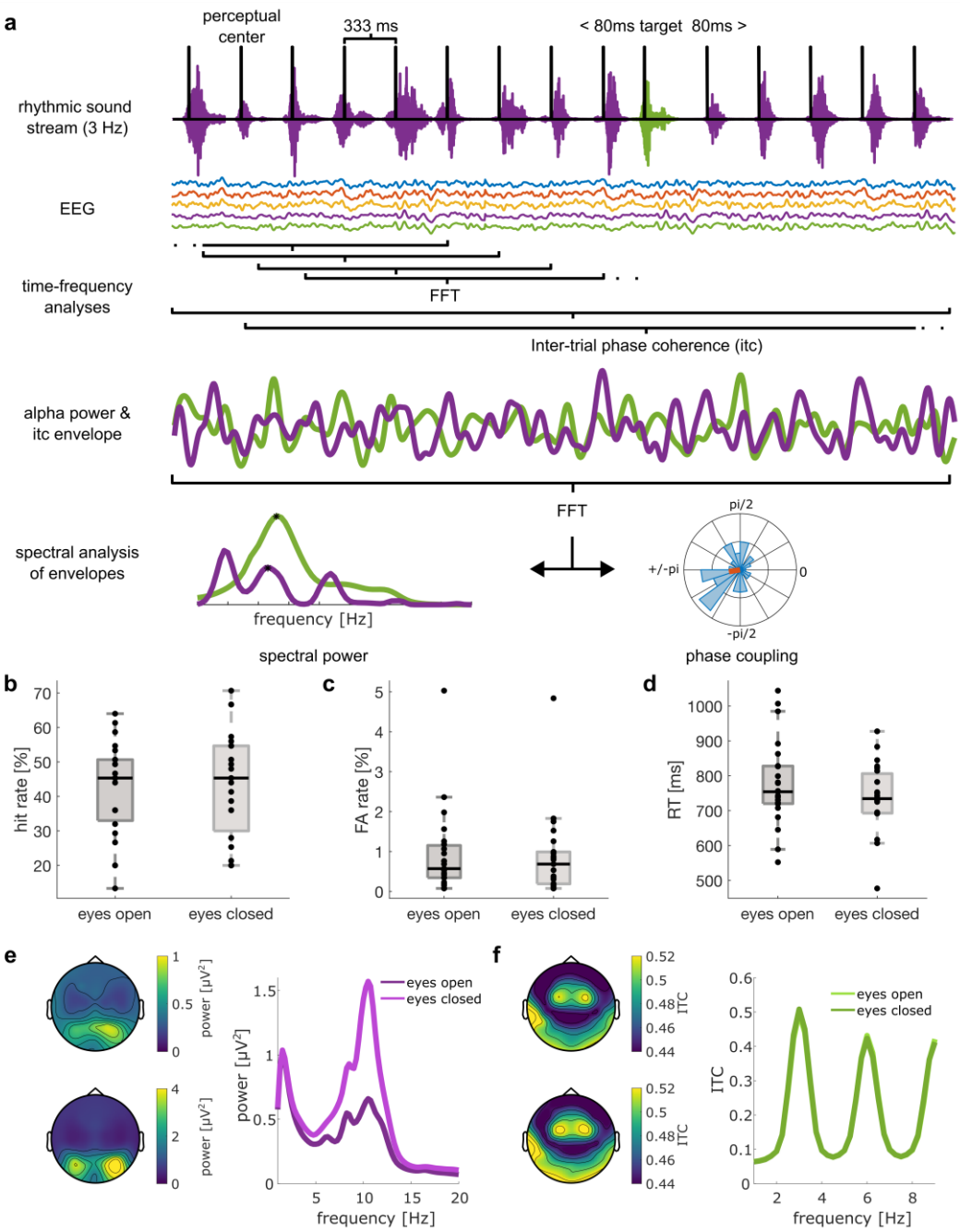



**Figure 1: Experimental design & analysis.**

**(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-closed 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. Boxplots depict the median of the data ± interquartile range. Whiskers indicate the range of the data.

## Results

### Overview

We recorded EEG from n=23 participants while they listened to 5-min streams of rhythmic, mono-syllabic, 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[23,29] (**Fig 1e,f**). α-oscillations showed a dominant occipito-parietal topography with a prominent increase in power when participants closed their eyes[29] (**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[19,23,30] (**Fig. 1f**).

### 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).

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 attentional 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.

**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., consistent phase relation) between α-power and entrainment across subjects 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 ± π ($CI_{99}$ = 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**). We could not identify reliable differences in oscillatory features or other variables that can distinguish these 15 participants from the others (**Supplementary Fig. S2, Supplementary Note 1).**

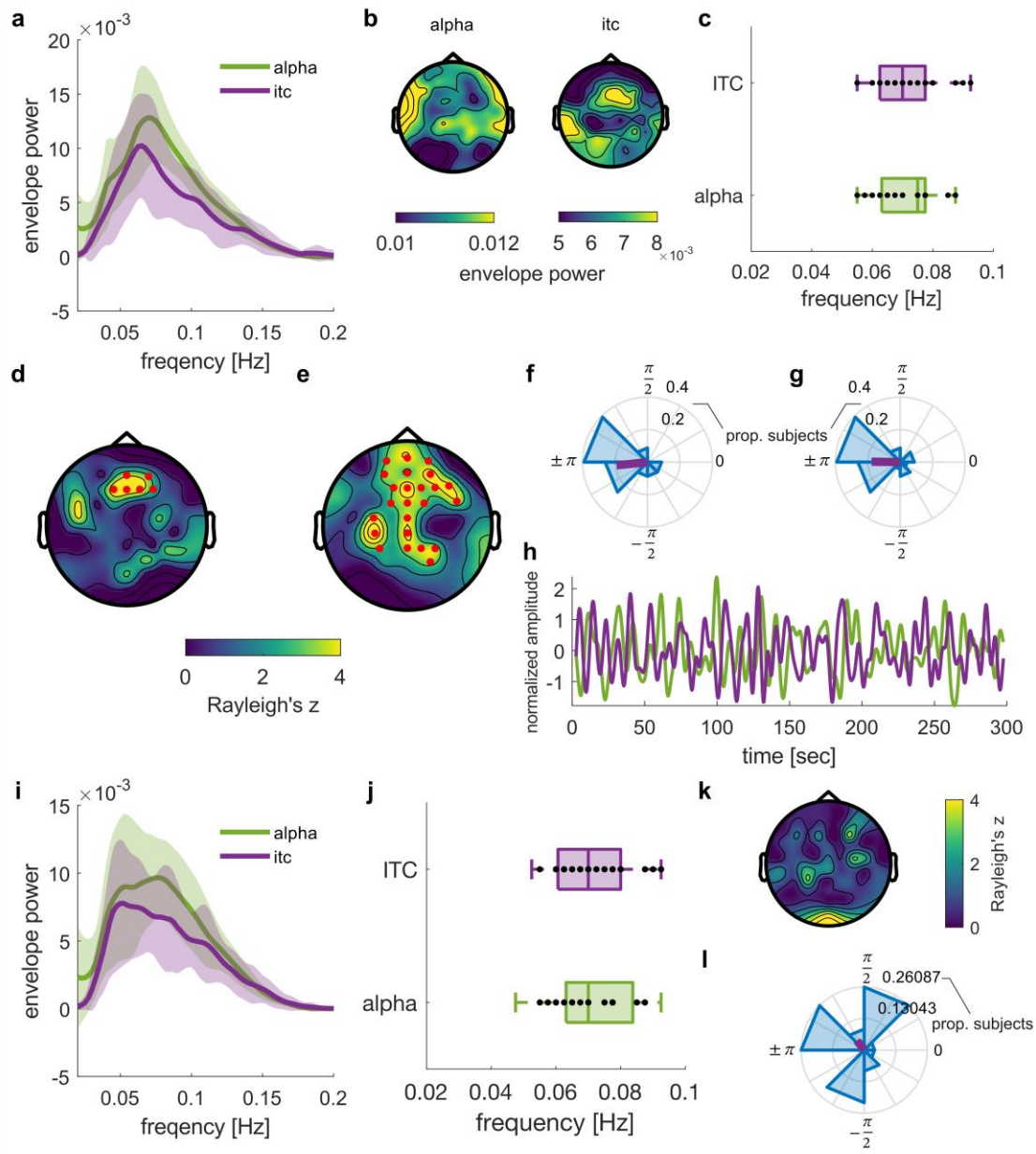



**Figure 2: α-power and entrainment exhibit slow anti-phase fluctuations.**

**(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). **(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 significant 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-closed condition. Shaded areas depict standard deviation. Boxplots depict the median of the data ± interquartile range. Whiskers indicate the range of the data.

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 ± π ($CI_{99}$ = 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. S3**). 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].

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

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.

**Slow rhythmic fluctuations cannot be explained by stimulus properties**

Although time intervals between targets were selected from a wide range and therefore irregular, their average (25 targets in 300 seconds) resembled the period of the slow neural fluctuations reported. To rule out that some regularity in target presentation explains the slow rhythmicity in α-

power or auditory entrainment, we computed the distribution of time intervals between targets (Fig. 3a), hits (Fig. 3b), and misses (Fig. 3c). None of these distributions shows a bias for 14 s (which would correspond to the 0.07 Hz rhythm found in α-power and ITC). Moreover, we found that individual frequencies for slow changes in α-power and ITC are uncorrelated with corresponding intervals between targets (Fig. 3d,g), hits (Fig. 3e,h), or misses (Fig. 3f,i). It therefore seems unlikely that the rate of target presentation gave rise to the observed fluctuations in the EEG.

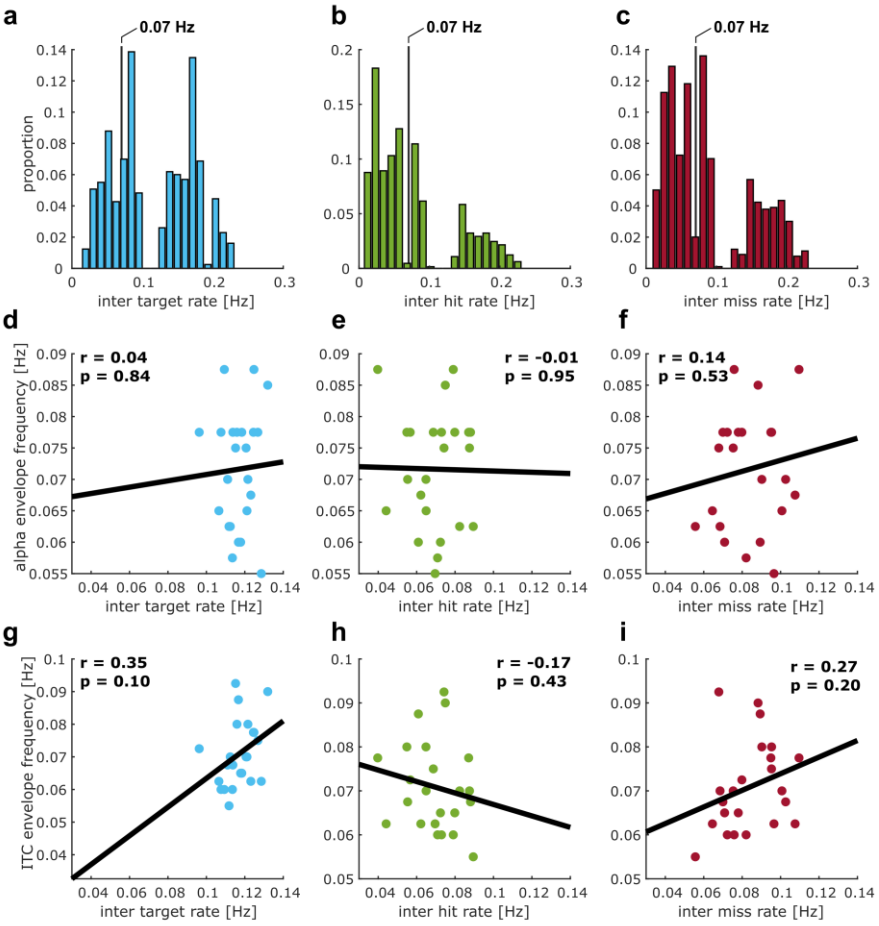



**Figure 3: Control analysis.**

(**a-c**) Distribution of time intervals between targets (**a**), hits (**b**), and misses (**c**), pooled across subjects for the eyes-open condition. Intervals were converted to rate (in Hz) for comparability with neural fluctuations. Vertical line indicates 0.07 Hz. (**d-i**) Correlations between individual frequencies for slow fluctuations in alpha power (middle row) or ITC (bottom row) and average intervals between targets (**d,g**), hits (**e,h**), or misses (**f,i**) while these fluctuations were measured.

**Opposing attentional modes emerge from interactions of distinct cortical networks**

To reveal the neural sources of opposing attentional modes, we source localized the slow rhythmic fluctuations in neural entrainment (**Fig. 4a**) and α-power (**Fig. 4b**) in the eyes-open condition, using a frequency domain beamformer[31] at the individual peak frequency of α-power fluctuations used in the previous analyses (**Fig. 2c**). For both of these, we observed activity in regions associated with auditory and speech processing including Superior Temporal Gyrus (STG), as well as left Pre-Central Gyrus. For entrainment **(Fig. 4a**), we found an additional involvement of left Inferior Frontal Gyrus (IFG) as well as the right Inferior Parietal Lobe, Angular Gyrus and parts of posterior STG. For α-power (**Fig. 4b**), we found additional activity in left Post-Central Gyrus, Superior Parietal Lobe as well as Posterior Cingulate and Occipital Cortex.

We then quantified coupling between neural entrainment and α-power as a phase relation between the two that is consistent across participants. This approach was similar to the previous sensor-level analysis, only that is was applied to different combinations of 360 cortical regions-of-interest (ROIs), parcellated on the source level according to the HCP-MMP1 atlas of the human connectome project[32]. We found significant coupling between slow entrainment changes bilaterally in ROIs in temporal cortex and left Inferior Frontal Gyrus, and equivalent changes in α-power in bilateral Posterior Cingulate Cortex, Superior Parietal Lobe and right Inferior Frontal Gyrus (permutation cluster Rayleigh-test with three significant clusters: $p_{cluster1} = .0026$, $p_{cluster2} = .004$, $p_{cluster3} = .006$; **Fig. 4c**). **Table 1** provides an overview of ROIs within the significant clusters, **Fig. 4d** depicts coupling for ROIs sorted into 44 higher order cortical structures according to the HCP-MMP1 atlas. Again, the coupling was driven by a phase relation between ROIs that was close to anti-phase (**Fig. 4e**). Whereas structures showing coupled fluctuations in entrainment are typically associated with processing of speech[33–36] (and therefore may reflect attention to sensory input), those with corresponding α-fluctuations involve regions associated with fronto-parietal attention networks[37,38], but also the default mode network[38,39]. Together, these results imply that fluctuations in opposing attentional modes may stem from an interaction between distinct cortical networks, one associated with active processing of external stimuli (i.e., the speech network), and the other concerned with up- and down-regulation of processing resources via inhibitory α-oscillations.

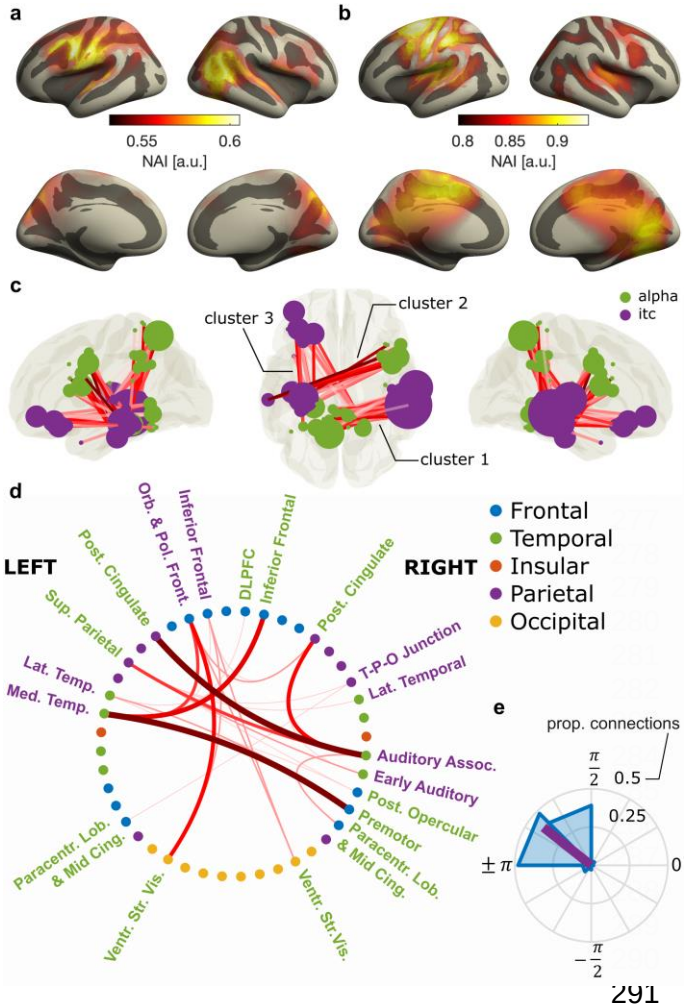



**Figure 4: Source-level analyses of α-power and entrainment fluctuations.**

**(a & b)** Neural activity index (NAI) of rhythmic fluctuations in entrainment (a) and α-power (b). **(c)** Coupling between α-power and entrainment fluctuations across 360 cortical ROIs. Each circle indicates the location of an ROI that belongs to a cluster with significant coupling. The size of the circle indicates the number of ROIs significantly coupled to each ROI. Lines indicate coupling between an ROI's entrainment fluctuations and another ROI's α-power fluctuations. Darker shades of red indicate stronger coupling. We found three clusters with significant coupling – 1: parietal (α) to right temporal (entrainment), 2: right frontal (α) to left temporal (entrainment), 3: parietal (α) to left frontal(entrainment) **(d)** Overview of cortical structures, extracted from the HCP-MMP1 atlas, that contain ROIs with coupling between α-power and entrainment fluctuations. Thicker, darker lines indicate more connections between ROIs in two given structures. Font color indicates if a structure shows (predominantly) entrainment (violet) or α-power coupling. **(e)** Distribution of phase differences between α-power and entrainment fluctuations from all ROIs that are part of the significant clusters.

## 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. 5a**,**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 ± π (**Fig. 5b,e)**, indicating an anti-phase relation**.** The observed phase relations appeared to be relatively stable across the pre-stimulus time period (**Fig. 5c,f**). Together, we found that the hypothesized markers of attentional processing (entrainment vs. α-power) differ depending on whether an auditory target was detected or not, as expected from modes of external and internal attention, respectively.

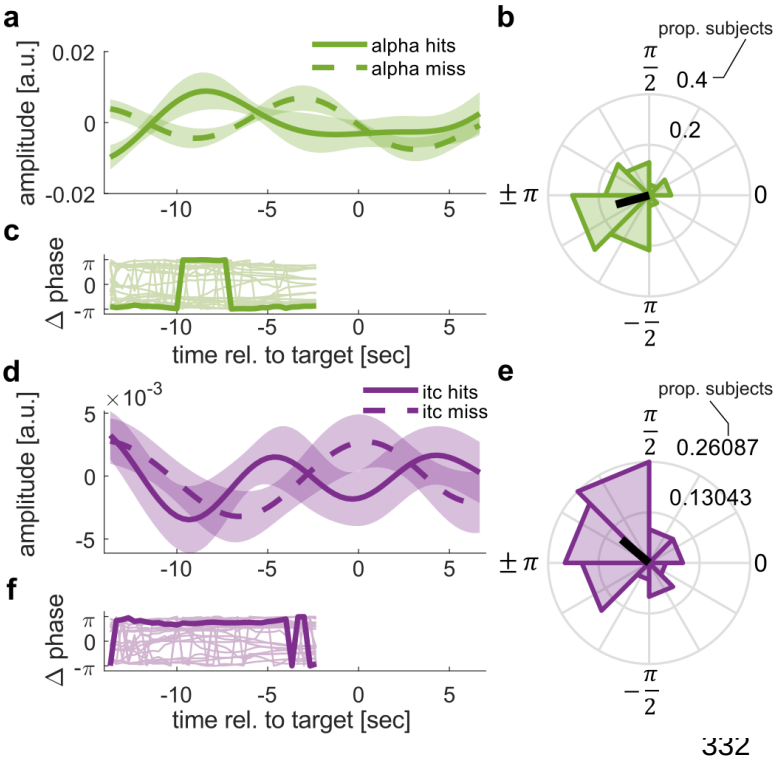



**Figure 5: α-power and entrainment fluctuations differ between detected and missed targets.**

(**a**) Average, bandpass filtered *α*-power fluctuations around hits and misses extracted from significant cluster shown in Fig. 2e. 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. Shaded areas indicate standard error of the mean.

## Discussion

Variations in 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. 5**). 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. The questions whether attention to sensory input always includes regular lapses, and which experimental manipulations can make attention more “sustained” need to be addressed in future work.

Our results indicate that fluctuations in attentional modes may emerge from interactions of distinct brain networks. Defined as neural activity aligned to a stimulus rhythm[19,20,22], we here used neural entrainment as a marker for sensory processing, and therefore external attention. We used rhythmic speech as the entraining stimulus, given its role in human communication and ability to entrain endogenous brain oscillations[19,20,22,23,40] In line with this approach, the network exhibiting slow, rhythmic fluctuations in entrainment entailed auditory and inferior frontal regions that typically entrain to auditory stimuli and speech in particular[40,41], and are involved in phonological and semantic processing[42].

Our results imply that, in moments with reduced external attention and corresponding decrease in neural entrainment, α-power increases in a network including Superior Parietal Lobe, Posterior Cingulate Cortex, and right IFG. Many of these regions have been associated with dorsal (and ventral) frontro-parietal attention networks[37,38,43]. These networks have previously been suggested to support attention in different modalities, although the exact regions involved may not overlap completely[44,45]. However, they have been predominantly studied in vision and evidence for their involvement in auditory, and particularly in sustained attention is scarce. Parts of Posterior Cingulate Cortex are also implicated in the default mode network[39]. Activity in this network has been associated with internal attention and reduced processing of external stimuli, resulting in attentional lapses[1,2]. Importantly, attention and default mode networks are typically localized with functional magnetic resonance imaging (fMRI), which offers high spatial resolution, but cannot resolve the role of brain oscillations in the regions. In contrast, previous research has linked α-oscillations to inhibition of sensory processing[13,14,46,47] and attentional (de-)selection[15–18,48]. In the context of sustained attention, high levels of α-oscillations are associated with attentional lapses[10,11,49] and periods of mind wandering[50]. However, these fields of research mostly exist in isolation, and it remained unclear to what extent brain networks found in fMRI and those that give rise to α-oscillation effects in attention overlap. Although

causality needs to be tested in follow-up work, our results suggest that dorsal attention and default mode networks may exercise rhythmic top-down control of sensory processing via α-oscillations, leading to slow, regular changes between internal and external attention.

While previous research has investigated the role of α-oscillations and neural entrainment to speech, only few works have considered them in conjunction[24–26]. Although two of these studies indicate some link between α-oscillations and neural entrainment[24,26], other work suggested that they are independent measures of neural activity[25], in apparent conflict with the current findings. However, most of this earlier work focused on fluctuations in α-power lateralization over time and consequently, on spatial rather than sustained attention. The "lateralization index" employed in these studies might have removed slow fluctuations in α-power that occur during sustained attention. Moreover, the trial-based experimental design used is - due to frequent interruptions or changes in stimulus presentation – likely inadequate to capture slow changes in attention that we hypothesized to occur during repetitive sensory stimulation. Accordingly, the constant 3-Hz rhythm in our experimental design might have been crucial to discover the slow anti-phasic fluctuations reported here. Indeed, research in the visual domain that did employ continuous reaction tasks[11,46,47] used α-power to predict lapses in attention up to 20 seconds before they occur. Interestingly, although not explored by the authors, α-power time courses seemed to contain a slow rhythmic component similar to the one observed here (cf. Fig. 5 in Ref[11]). Nevertheless, the fact that α-oscillations and neural entrainment fluctuate in slow anti-synchrony during sustained attention but not necessarily on a shorter time scale is an interesting observation that deserves closer examination.

It remains unclear why the observed effects vanish when participants close their eyes. Closing the eyes is known to cause a substantial increase of α-oscillations pre-dominantly in visual areas[29]. 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[51], while the fluctuations we observe in our data vary across participants. In a control analysis, we did not find any relationship between the rate of targets and the frequency of slow α-power and ITC fluctuations (**Fig. 3**).

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[52–59]. It may thus be plausible that rhythmicity in attention can also exists at other time scales. We speculate that regular changes between internal and external attention reflect a protective mechanism to prevent depletion of attentional resources. Fluctuations in attention may allow the system to maintain a higher-level of performance for a longer period of time at the cost of regular periods of reduced sensory processing. Indeed, evidence from non-invasive brain stimulation suggests that enhancing endogenous α-oscillations with electrical stimulation has a stabilizing, rather than a decremental effect on sustained attention[60].

There is an intriguing similarity of timescales between the intrinsic attentional rhythms in our data and the known coupling between α-oscillations and slow rhythmic activity in the respiratory system[61], the heart, and the gut[62]. In particular, the gastric network seems to generate rhythmic activity at similar time scales as those we found here (~0.05 Hz)[62,63], and modulate the amplitude of spontaneous α-oscillations[64]. Future research should address if these similarities in time scales are functionally meaningful. A recent magnetoencephalography (MEG) study supports this notion as it localized gastric-alpha coupling to right IFG and parietal lobe[64], regions that overlap with the attention-related networks reported in our study.

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[65–67] and has been previously shown to stabilize sustained attention when applied in a continuous manner[60]. 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]. Our source-localization results reveal regions that can be targeted to modulate attentional modes. 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[68]. It thus remains to be determined if similar rhythmicity exists when sustained attention is deployed to visual, somatosensory, or cross-modal tasks and if so, at which frequencies rhythmicity emerges, and whether the same attentional networks engage in its regulation. First evidence for similar effects in the visual domain has already been reported in non-human primates[10]

## Materials & Methods

### 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 (protocol no: CPP 21.01.22.71950 / 2021-A00131-40). All ethical regulations relevant to human research participants were followed.

### 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.

### 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)[69] of the words to the metronome beat and resulted in perceptually very rhythmic speech (see Zoefel et al[70] 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[71]. 16-channel noise-vocoded speech is not as easy to understand as clear speech, but is still clearly intelligible[72]. Individual, noise-vocoded words were then concatenated into a continuous sound stream of 5-min length (i.e.,

totaling 900 words). The order of words was randomized with a constraint such that every word from the stimulus set had occurred before it can be repeated. Across all blocks, 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 (25 targets per block, 75 targets per eyes-open/eyes-closed condition). Participants were asked to indicate violations of the rhythm by pressing the space bar on a standard computer keyboard.

The experiment was carried out in a dimly-lit recording chamber, separated from the experimenter. 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 a pneumatic In-Ear headphone system (Etymotic Research ER-2, Etymotic Research Inc., USA). This system provides an additional layer of shielding from environmental noise. 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 on a black background at the center of the screen. The fixation cross was also shown during eyes-closed conditions to keep environmental light conditions constant across blocks.

**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.

**EEG processing**

EEG analyses were performed in MATLAB 2019b using the fieldtrip toolbox[73]. 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. Power and ITC spectra were visually inspected to ensure good ratio between spectral peaks and the "noise" at the surrounding frequencies as well as to ensure plausible topographies.

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[74] 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. S3**.

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 α-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

552 values and to smoothen α-power trajectories. The analysis was restricted to channels from signifi-
553 cant clusters revealed in previous analyses (**Fig. 2d,e**; **Supplementary Table 1**).

554 **Source analysis**

555 We applied dynamic imaging of coherent sources (DICS) beamforming[31] at individual peak frequen-
556 cies of the slow fluctuations in α-power and ITC. Leadfields were constructed from the standard
557 boundary element model of the fieldtrip toolbox, and standardized electrode locations were defined
558 in MNI space. We projected the complex Fourier coefficients from each 100-sec segment onto a
559 surface grid with 20484 source locations. From these complex coefficients, we obtained a neural
560 activity index by dividing oscillatory power by an estimate of the noise bias. To quantify coupling
561 between α-power and ITC fluctuations for different combinations of brain regions, we computed the
562 phase angle of these two neural measures for each segment, source location, and subject. We then
563 applied a parcellation into 360 regions of interest (ROIs) according to the HCP-MMP1 atlas of the
564 human connectome project[32]. To this end, we computed the average phase across all source loca-
565 tions within an ROI. We then computed the average phase across the 60 segments for each ROI
566 and subject, separately for α-power and ITC fluctuations, and tested for consistent phase relations
567 between the two (our measure of connectivity) across subjects, as detailed below. The parcellation
568 of source-level data reduces the number of channel-channel comparisons necessary and provides
569 meaningful anatomical labels, thereby improving interpretability and reducing computational de-
570 mands. It also mitigates localization errors arising from the use of standard head models and elec-
571 trode locations for all participants, as brain activity is integrated over larger areas of the brain.

572 **Behavioral analysis**

573 Average reaction times, as well as the proportion of hits, misses and false alarms were computed
574 for each condition (eyes-open vs. eyes-closed). A target word was considered a hit, if a button was
575 pressed within 2-sec after stimulus onset, otherwise it was considered a miss. Button presses oc-
576 curring outside these response intervals were considered false-alarms. Hit rates were computed by
577 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).

**Statistical analyses & Reproducibility**

Statistical analyses were performed in Matlab 2019b using the circular statistics toolbox[75] in combination with functions in fieldtrip for massive-multivariate permutation statistics[73,76]. All statistical analysis were conducted on n=23 participants.

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[76].

A similar statistical approach was used to test for significant coupling of α-power and ITC fluctuations in the parcellated source-level data. In contrast to the sensor level analysis, the test was run for all possible combinations of ROIs. To keep the resulting network structures sparse, we employed a more conservative α-level of 0.01 for this test.

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 α-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 α-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 α-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 $\pm\pi$.

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

## Data availability

The underlying anonymized data is available in minimally processed form (down sampled, filtered, artifacts suppressed with ica) via the open science framework (https://osf.io/64ycj/). Due to the file size, the unprocessed raw data can be obtained upon reasonable request from the corresponding authors.

## Code availability

The MATLAB code used to process and analyze the data is available via the same open science framework repository (https://osf.io/64ycj/).

## 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.

## 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

## Conflict of interest

FHK, QB, and BZ declare no competing interests.

## 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.

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

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

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

677 17. Okazaki, Y.O., De Weerd, P., Haegens, S., and Jensen, O. (2014). Hemispheric lateralization of
678 posterior alpha reduces distracter interference during face matching. Brain Research *1590*, 56–
679 64. 10.1016/j.brainres.2014.09.058.

680 18. Wöstmann, M., Herrmann, B., Maess, B., and Obleser, J. (2016). Spatiotemporal dynamics of
681 auditory attention synchronize with speech. Proc Natl Acad Sci USA *113*, 3873–3878.
682 10.1073/pnas.1523357113.

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

685 20. Henry, M.J., and Obleser, J. (2012). Frequency modulation entrains slow neural oscillations and
686 optimizes human listening behavior. Proceedings of the National Academy of Sciences *109*,
687 20095–20100. 10.1073/pnas.1213390109.

688 21. Riecke, L., Formisano, E., Sorger, B., Başkent, D., and Gaudrain, E. (2018). Neural Entrainment
689 to Speech Modulates Speech Intelligibility. Current Biology *28*, 161-169.e5.
690 10.1016/j.cub.2017.11.033.

691 22. Henry, M.J., Herrmann, B., and Obleser, J. (2014). Entrained neural oscillations in multiple fre-
692 quency bands comodulate behavior. Proceedings of the National Academy of Sciences *111*,
693 14935–14940. 10.1073/pnas.1408741111.

694 23. van Bree, S., Sohoglu, E., Davis, M.H., and Zoefel, B. (2021). Sustained neural rhythms reveal
695 endogenous oscillations supporting speech perception. PLOS Biology *19*, e3001142.
696 10.1371/journal.pbio.3001142.

697 24. Kerlin, J.R., Shahin, A.J., and Miller, L.M. (2010). Attentional Gain Control of Ongoing Cortical
698 Speech Representations in a “Cocktail Party.” J. Neurosci. *30*, 620–628. 10.1523/JNEURO-
699 SCI.3631-09.2010.

700 25. Tune, S., Alavash, M., Fiedler, L., and Obleser, J. (2021). Neural attentional-filter mechanisms
701 of listening success in middle-aged and older individuals. Nat Commun *12*, 4533.
702 10.1038/s41467-021-24771-9.

703 26. Hauswald, A., Keitel, A., Chen, Y.-P., Rösch, S., and Weisz, N. (2022). Degradation levels of
704 continuous speech affect neural speech tracking and alpha power differently. European Journal
705 of Neuroscience *55*, 3288–3302. 10.1111/ejn.14912.

706 27. Pöppel, E. (1997). A hierarchical model of temporal perception. Trends in Cognitive Sciences *1*,
707 56–61. 10.1016/S1364-6613(97)01008-5.

708 28. Ding, N., Patel, A.D., Chen, L., Butler, H., Luo, C., and Poeppel, D. (2017). Temporal modula-
709 tions in speech and music. Neurosci Biobehav Rev *81*, 181–187. 10.1016/j.neubio-
710 rev.2017.02.011.

29. 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.

30. 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.

31. Gross, J., Kujala, J., Hamalainen, M., Timmermann, L., Schnitzler, A., and Salmelin, R. (2001). Dynamic imaging of coherent sources: Studying neural interactions in the human brain. Proceedings of the National Academy of Sciences *98*, 694–699. 10.1073/pnas.98.2.694.

32. Glasser, M.F., Coalson, T.S., Robinson, E.C., Hacker, C.D., Harwell, J., Yacoub, E., Ugurbil, K., Andersson, J., Beckmann, C.F., Jenkinson, M., et al. (2016). A multi-modal parcellation of human cerebral cortex. Nature *536*, 171–178. 10.1038/nature18933.

33. Davis, M.H., and Johnsrude, I.S. (2003). Hierarchical Processing in Spoken Language Comprehension. J. Neurosci. *23*, 3423–3431. 10.1523/JNEUROSCI.23-08-03423.2003.

34. Hickok, G., and Poeppel, D. (2007). The cortical organization of speech processing. Nat Rev Neurosci *8*, 393–402. 10.1038/nrn2113.

35. Friederici, A.D., and Gierhan, S.M. (2013). The language network. Current Opinion in Neurobiology *23*, 250–254. 10.1016/j.conb.2012.10.002.

36. Schroën, J.A.M., Gunter, T.C., Numssen, O., Kroczek, L.O.H., Hartwigsen, G., and Friederici, A.D. (2023). Causal evidence for a coordinated temporal interplay within the language network. Proceedings of the National Academy of Sciences *120*, e2306279120. 10.1073/pnas.2306279120.

37. Corbetta, M., and Shulman, G.L. (2011). SPATIAL NEGLECT AND ATTENTION NETWORKS. Annu Rev Neurosci *34*, 569–599. 10.1146/annurev-neuro-061010-113731.

38. Uddin, L.Q., Yeo, B.T.T., and Spreng, R.N. (2019). Towards a Universal Taxonomy of Macroscale Functional Human Brain Networks. Brain Topogr *32*, 926–942. 10.1007/s10548-019-00744-6.

39. Smallwood, J., Bernhardt, B.C., Leech, R., Bzdok, D., Jefferies, E., and Margulies, D.S. (2021). The default mode network in cognition: a topographical perspective. Nat Rev Neurosci *22*, 503–513. 10.1038/s41583-021-00474-4.

40. Kösem, A., Bosker, H.R., Takashima, A., Meyer, A., Jensen, O., and Hagoort, P. (2018). Neural Entrainment Determines the Words We Hear. Current Biology *28*, 2867-2875.e3. 10.1016/j.cub.2018.07.023.

41. Peelle, J.E., Gross, J., and Davis, M.H. (2013). Phase-Locked Responses to Speech in Human Auditory Cortex are Enhanced During Comprehension. Cerebral Cortex *23*, 1378–1387. 10.1093/cercor/bhs118.

42. Hodgson, V.J., Lambon Ralph, M.A., and Jackson, R.L. (2021). Multiple dimensions underlying the functional organization of the language network. NeuroImage *241*, 118444. 10.1016/j.neuroimage.2021.118444.

43. Thakral, P.P., and Slotnick, S.D. (2009). The role of parietal cortex during sustained visual spatial attention. Brain Research *1302*, 157–166. 10.1016/j.brainres.2009.09.031.

44. Braga, R.M., Wilson, L.R., Sharp, D.J., Wise, R.J.S., and Leech, R. (2013). Separable networks for top-down attention to auditory non-spatial and visuospatial modalities. NeuroImage *74*, 77–86. 10.1016/j.neuroimage.2013.02.023.

45. Tobyne, S.M., Osher, D.E., Michalka, S.W., and Somers, D.C. (2017). Sensory-biased attention networks in human lateral frontal cortex revealed by intrinsic functional connectivity. NeuroImage *162*, 362–372. 10.1016/j.neuroimage.2017.08.020.

46. 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.

47. 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.

48. 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.

49. 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.

50. 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.

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

52. 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.

53. 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.

54. 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.

55. 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.

56. 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.

57. 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.

58. 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.

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

60. 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.

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

62. 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.

63. 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.

64. 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.

65. 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.

66. 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.

67. 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.

68. 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.

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

70. 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.

71. 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.

72. 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.

73. 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.

74. 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.

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

76. 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.

# Tables

**Table 1**

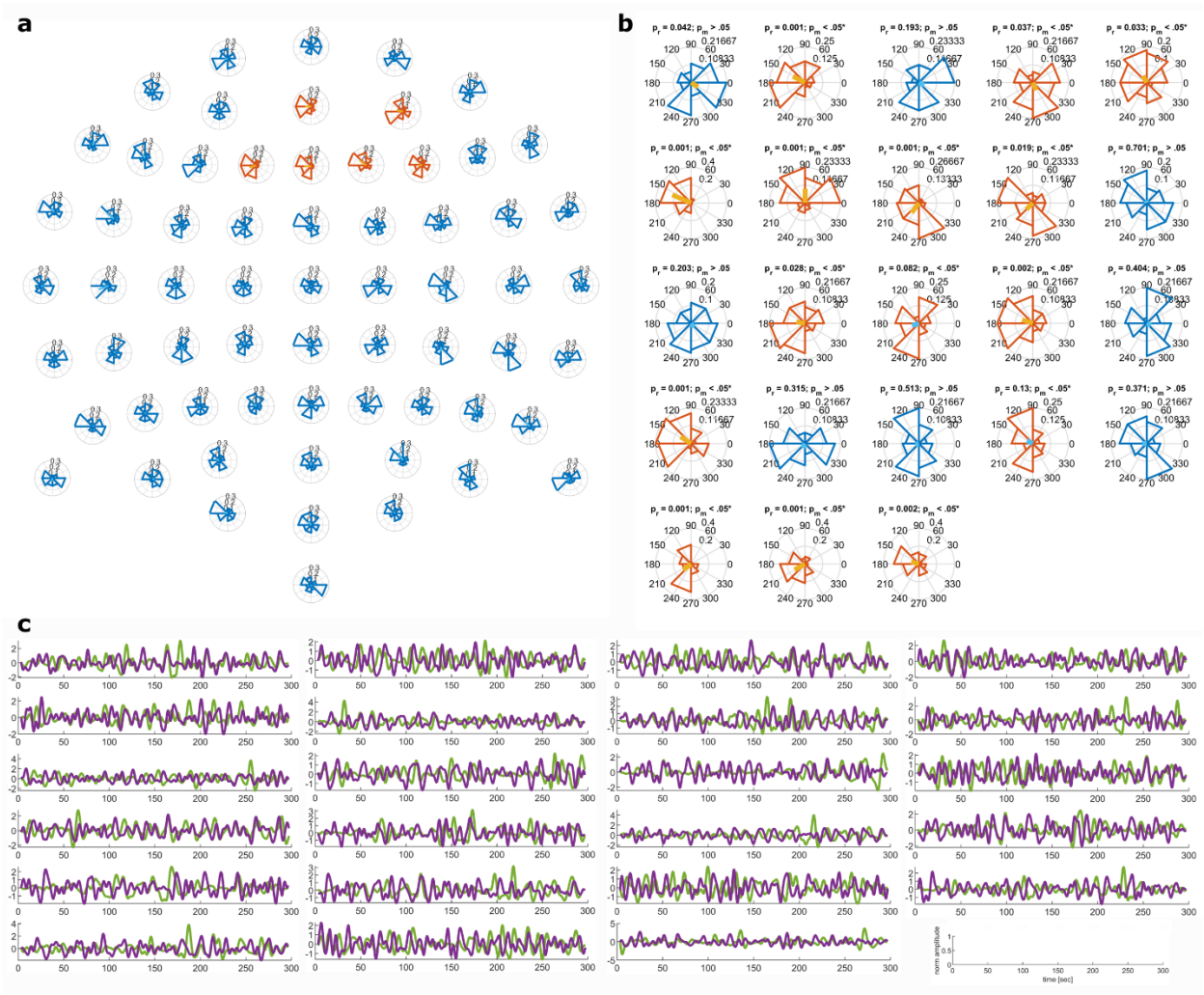
*Regions according to HCP-MMP1 atlas showing auditory entrainment or α-power coupling.*

| Cortical structure | Measure | ROIs | Hemisphere |
|---|---|---|---|
| Medial temporal | ITC | Hippocampus, Para-Hippocampal Area 3, Area TF, Para-Hippocampal Area 2 | L |
| Lateral Temporal | ITC | Area TE1 Middle, Area TG dorsal | L |
| Inferior Frontal | ITC | Area 47l (47 lateral), Area anterior 47r, | L |
| Orbital and Polar Frontal | ITC | Area 13l, Area 47s, Area 47m | L |
| Posterior Opercular | ITC | Area OP4-PV | R |
| Early Auditory | ITC | Para-Belt Complex, Medial Belt Complex | R |
| Auditory Association | ITC | Area STSd posterior, Area STSv posterior, Auditory 4 Complex | R |
| Temporo-Parieto-Occipital Junction | ITC | Area Temporo-Parieto-Occipital Junction 1 | R |
| Lateral Temporal | ITC | Area TE1 Middle | R |
| Posterior Cingulate | alpha | Retro-Splenial Complex, Pre-Cuneus Visual Area, Area dorsal 23 a+b, Area 31p ventral, Area 31pd, Area 31a | L |
| Paracentral Lobular & Mid Cingulate | alpha | Area 5L | L |
| Superior Parietal | alpha | Lateral Area 7A, Medial Area 7A | L |
| Ventral Stream Visual | alpha | Fusiform Face Complex, Ventro-Medial Visual Area 1, Ventral Visual Complex | L |
| Medial Temporal | alpha | Pre-Subiculum, Para-Hippocampal Area 1 | L |
| Premotor | alpha | Premotor Eye Field, Area 55b, Rostral Area 6 | R |
| Posterior Cingulate | alpha | Pre-Cuneus Visual Area, Area dorsal 23 a+b, Area 31a, Pro-Striate Area | R |
| Paracentral Lobular & Mid Cingulate | alpha | Area 5m, Area 5L | R |
| Dorsolateral Prefrontal | alpha | Area 8C | R |
| Inferior Frontal | alpha | Area 44, Area IFJa | R |
| Posterior Opercular | alpha | Area OP2-3-VS | R |
| Ventral Stream Visual | alpha | Ventro-Medial Visual Area 1 | R |

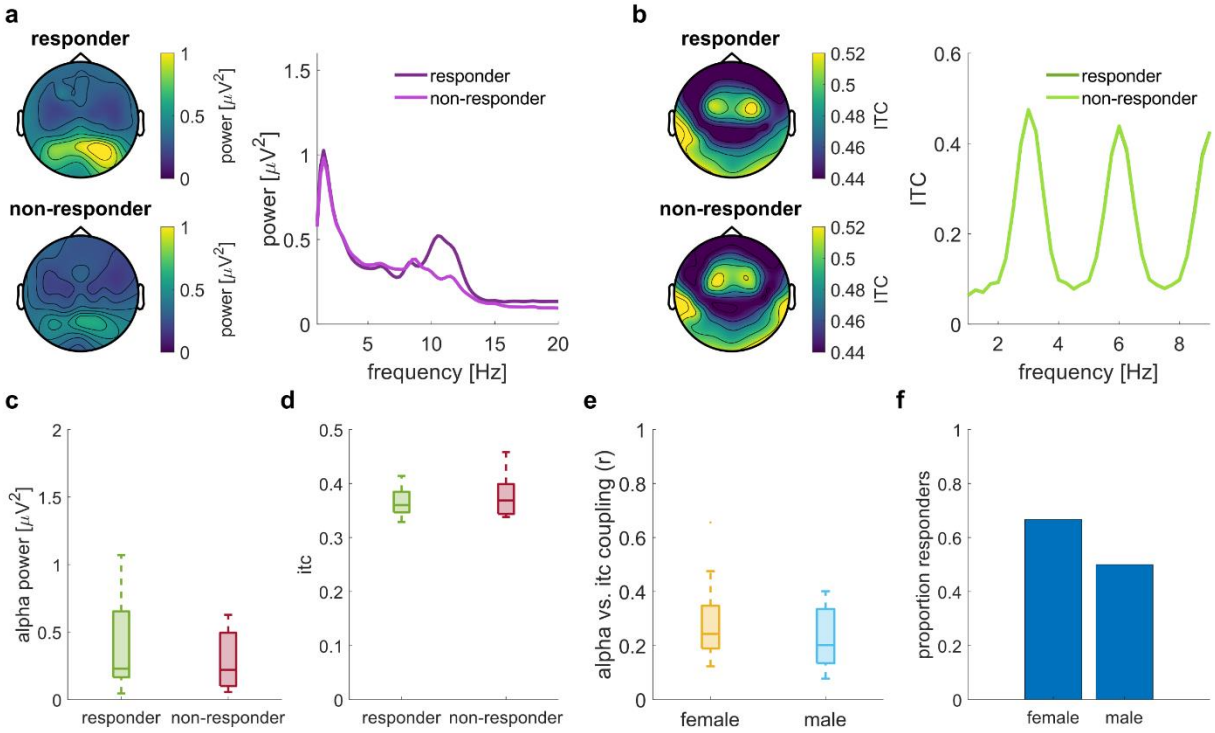
*Note.* Overview of cortical structures with ROIs that show auditory entrainment or α-power fluctuations significantly coupled to other ROIs. Labels are according to the HCP-MMP1 atlas**.** *Measure* indicates if ROIs show ITC coupled to α-power in other ROIs (*ITC*) or α-power coupled to ITC in other ROIs (*alpha*). Hemisphere labels correspond to: L = left, R = right.

# Supplementary Information: Opposing neural processing modes during auditory sustained attention alternate rhythmically

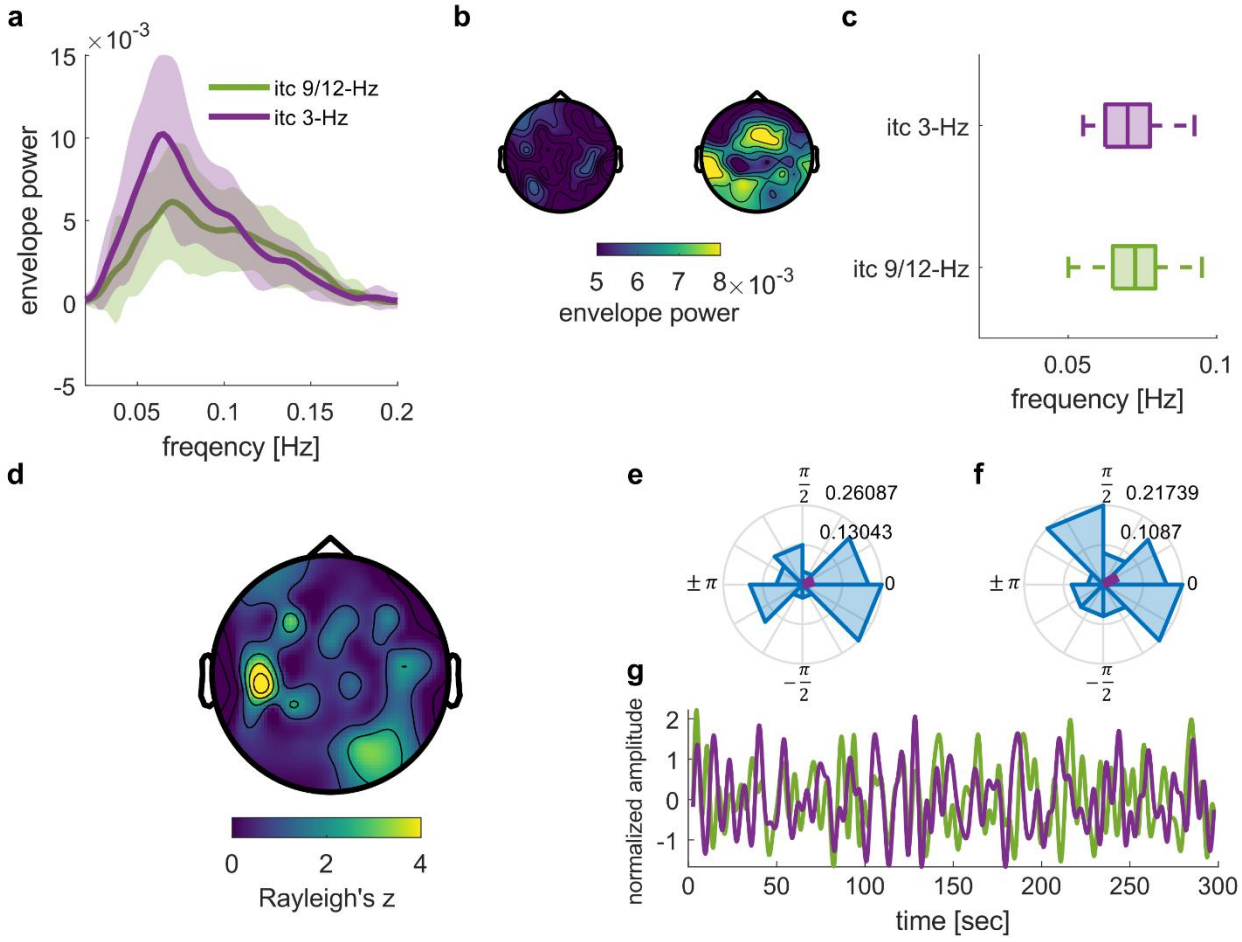
# 1 Supplementary Figures




**Supplementary Figure S6: α-power and entrainment coupling for individual channels and participants.** **(a + b)** Distributions of phase differences between α-power and entrainment fluctuations. (a) shows these distributions across participants but separately for each channel. Distributions in orange indicate channels forming the significant cluster revealed by the permutation Rayleigh test (Fig. 2d in the main manuscript). Anti-phase coupling appears to be consistent across channels within the cluster. (**b**) shows distributions for individual subjects, within the significant channel cluster. Above the individual panels are shown results from single-subject Rayleigh's test ($p_r$) for non-uniformity computed across 60 overlapping 100-sec segments, and from test comparing the trial-averaged angle against 0 ($p_m$). 14 out of 23 participants exhibit a significant coupling between α-power and entrainment fluctuations at an α-level of 0.05. 15 out of 23 participants further exhibit an average phase difference that is significantly different from 0. Phase distributions with significant coupling are indicated by a yellow average phase vector. Distributions significantly differing from zero are shown in orange. Importantly, there is a strong overlap between subjects showing significant coupling of the signals, and those showing an average phase difference significantly different from zero. (**c**) Exemplary α-power and entrainment time-courses. Traces depict fluctuations of α-power (green) and auditory entrainment (violet) over time for an exemplary experimental block with eyes-open for each participant. For display purposes, the traces were filtered between 0.03-Hz and 0.2-Hz. Axis labels are indicated in the lower right plot.




**Supplementary Figure S7: Responder/Non-responder analysis (eyes-open condition).** (**a**) Topography of average α-power for responders (top left) and non-responders (bottom left). The power spectrum on the right depicts the average power spectra of responders and non-responders within the cluster of electrodes showing significant α-power and ITC coupling (Fig. 2e in main manuscript). (**b**) Topography of 3-Hz ITC for responders (top left) and non-responders (bottom left). The ITC spectrum on the right depicts the average ITC across frequencies within the cluster of electrodes showing significant α-power and ITC coupling (Fig. 2d in the main manuscript). (**c**) Boxplots depict the distribution of α-power within the cluster of electrodes showing significant coupling to ITC for responders and non-responders. Panel (**d**) depicts the same distribution for ITC values. (**e**) Distribution of α-power and coupling (r, cf. Supplementary Fig. S1b), separately for male and female participants. (**f**) Proportion of male and female participants showing significant α-power and ITC coupling.




**Supplementary Figure S8: Control analysis for coupling between 3-Hz entrainment and entrainment at harmonic frequencies that lie in the α-band (9-Hz and 12-Hz).** To rule out that the coupling between α-band power and 3-Hz entrainment is driven by entrainment effects at harmonic frequencies within the α-band, we repeated the analysis described in the main manuscript. However, instead of using α-alpha power fluctuations, we tested whether 3-Hz ITC fluctuations are coupled with ITC fluctuations at harmonics of 3 Hz and that lie in the α-band (9-Hz and 12-Hz). (**a**) The resulting spectra show a substantially reduced rhythmic activity compared to the fluctuations in α-power (Fig. 2a). (**b**) Topography of fluctuations in 9-Hz and 12-Hz ITC (left) as well as 3-Hz ITC (right, corresponding to Fig. 2b in main manuscript). (**c**) Distribution of peak frequencies in the spectra of α-power and ITC envelopes. (**d**) Permutation cluster analysis does not reveal any significant coupling between 3-Hz and 9-Hz & 12Hz ITC envelopes ($p_{cluster} > .17$). (**e+f**) Distribution of phase differences between 3-Hz and 9-Hz &12-Hz ITC fluctuations, within the clusters that show significant coupling between α-power and 3-Hz ITC (Fig. 2d). We did not find evidence that these phases significantly differ from 0 (circular one-sample test against angle of zero: $p > .05$, $M_{angle} = 0.28$ rad, and $p > .05$, $M_{angle} = 0.49$ rad). (**g**) exemplary time-course of 3-Hz vs. 9-Hz + 12-Hz ITC fluctuations. Overall, these results do not support the idea that harmonic entrainment in the α-band gives rise to the coupling between α-power fluctuations and auditory entrainment.

## 2 Supplementary Tables

***Supplementary Table 1****: Composition of clusters with significant entrainment vs. alpha power coupling in permutation cluster analysis*

| Cluster | Channels | $p_{cluster}$ |
|---|---|---|
| ROI1: entrainment vs. alpha per channel (Fig. 2d) | F1, AF4, AFz, Fz, F2, F4 | .037 |
| ROI 2: entrainment in ROI 1 vs. alpha per channel (Fig. 2e) | Fp1, AF3, F1, FC1, C3, CP3, P3, Pz, CPz, Fpz, Fp2, AF4, AFz, Fz, F2, F4, F6, FC6, FC2, FCz, Cz, P2, P4, PO4 | .042 |

## 3 Supplementary Notes

### Supplementary Note 1

In an exploratory follow-up analysis, we contrasted various oscillatory features in EEG data from participants which exhibit coupling between α-power and entrainment fluctuations (15 “responders” in **Supplementary Fig. S1b**) with features from those who did not (8 “non-responders” in Supplementary Fig. S1b). Responders exhibited numerically higher power in the α-band (Supplementary Fig. S2**a,c**) despite near-identical ITC (**Supplementary Fig. S2b,d**). However, when we compared α-power within the significant cluster that showed coupling to ITC (cf. Fig. 2e), this difference did not reach significance (responder vs non-responder α-power: $t_{21} = 0.75$, $p = .46$; ITC: $t_{21} = -0.74$, $p = 0.47$; independent samples t-test). There was no reliable difference in α-power and ITC coupling between male and female participants ($t_{21} = 0.95$, $p = .35$, independent samples t-test; Supplementary Fig. S2e,f)