



HAL
open science

The inattentional rhythm in audition

Troby Ka-Yan Lui, Eva Boglietti, Benedikt Zoefel

► **To cite this version:**

Troby Ka-Yan Lui, Eva Boglietti, Benedikt Zoefel. The inattentional rhythm in audition. 2024.
hal-04758130

HAL Id: hal-04758130

<https://hal.science/hal-04758130v1>

Preprint submitted on 29 Oct 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27

The inattentional rhythm in audition

Troby Ka-Yan Lui*^{1,2}, Eva Boglietti¹, Benedikt Zoefel*^{1,2}

Université de Toulouse III Paul Sabatier, Toulouse, France

Centre National de la Recherche Scientifique (CNRS), Centre de Recherche Cerveau et
Cognition (CerCo), UMR 5549, Toulouse, France

* Corresponding Author: Troby Ka-Yan Lui (trobylui@gmail.com), Benedikt Zoefel
(benedikt.zoefel@cnrs.fr)

Acknowledgements: This study was supported by a grant from the Agence Nationale de la
Recherche (ANR-21-CE37-0002). The authors thank Quentin Busson for help with data collection.

28

Abstract

29 The detection of temporally unpredictable visual targets depends on the preceding phase of alpha
30 oscillations (~7-12 Hz). In audition, however, such an effect seemed to be absent. Due to the
31 transient nature of its input, the auditory system might be particularly vulnerable to information
32 loss that occurs if relevant information coincides with the low excitability phase of the oscillation.
33 We therefore hypothesised that effects of oscillatory phase in audition will be restored if auditory
34 events are made task-irrelevant and information loss can be tolerated. To this end, we collected
35 electroencephalography (EEG) data from 29 human participants (21F) while they detected pure
36 tones at one sound frequency and ignored others. Confirming our hypothesis, we found that the
37 neural response to task-irrelevant but not to task-relevant tones depends on the pre-stimulus phase
38 of neural oscillations. Alpha oscillations modulated early stages of stimulus processing, whereas
39 theta oscillations (~3-7 Hz) affected later components, possibly related to distractor inhibition. We
40 also found evidence that alpha oscillations alternate between sound frequencies during divided
41 attention. Together, our results suggest that the efficacy of auditory oscillations depends on the
42 context they operate in, and demonstrate how they can be employed in a system that heavily relies
43 on information unfolding over time.

44

45

Significance Statement

46 The phase of neural oscillations shapes visual processing, but such an effect seemed absent in the
47 auditory system when confronted with temporally unpredictable events. We here provide evidence
48 that oscillatory mechanisms in audition critically depend on the degree of possible information
49 loss during the oscillation's low excitability phase, possibly reflecting a mechanism to cope with
50 the rapid sensory dynamics that audition is normally exposed to. We reach this conclusion by
51 demonstrating that the processing of task-irrelevant but not task-relevant tones depends on the pre-
52 stimulus phase of neural oscillations during selective attention. During divided attention, cycles of
53 alpha oscillations seemed to alternate between possible acoustic targets similar to what was
54 observed in vision, suggesting an attentional process that generalises across modalities.

55

56

Introduction

57 Confronted with a dynamic environment, our brain constantly engages in the selection and
58 prioritization of incoming sensory information. Previous research posits that neural oscillations,
59 rhythmic fluctuations in neural excitability, are instrumental for this purpose (Schroeder &
60 Lakatos, 2009). One fundamental assumption in this line of research is that the sensory information
61 that coincides with the high-excitability phase of an oscillation is processed more readily than that
62 occurring during the low-excitability phase, leading to perceptual or attentional rhythms
63 (VanRullen, 2016b).

64 Previous studies in the visual modality have confirmed this assumption, demonstrating that
65 the detection of temporally unpredictable targets depends on the pre-stimulus phase of alpha
66 oscillations in the EEG (Busch et al., 2009; Dugué et al., 2015; Dugué et al., 2011; Mathewson et
67 al., 2009). This phasic effect was only found for the detection of attended, but not unattended
68 visual targets (Busch & VanRullen, 2010).

69 Studies in the auditory modality, however, revealed a more ambivalent role of neural
70 oscillations in auditory perception (VanRullen et al., 2014). On the one hand, the detection of near-
71 threshold auditory tones, presented at unpredictable moments in quiet, does not depend on pre-
72 target neural phase (VanRullen et al., 2014; Zoefel & Heil, 2013). This result seems to question
73 the assumption of an auditory perception that is inherently rhythmic. On the other hand, it is clear
74 that stimulus-aligned (“entrained”) neural oscillations serve a mechanistic role in auditory
75 attention and perception (Obleser & Kayser, 2019; Henry & Obleser, 2012; van Bree et al., 2021).
76 Rhythmicity in auditory processing can also be observed after a cue like the onset of acoustic noise,
77 assumed to reflect a phase reset of oscillations in the theta range (Ho et al., 2017; Lui et al., 2023;
78 Wöstmann et al., 2020).

79 We here tested a hypothesis that can reconcile these apparently discrepant findings. This
80 hypothesis is based on the fact that the auditory environment is particularly dynamic and transient
81 (Kubovy, 1988; VanRullen et al., 2014). Losing critical auditory information that coincides with
82 the low-excitability phase of the oscillation may be too costly for successful comprehension of
83 auditory input. To avoid such a loss of information, the brain may therefore suppress neural
84 oscillations in auditory system and operate in a more “continuous mode” (Schroeder & Lakatos,
85 2009) if incoming auditory stimuli are relevant (e.g., attended) but their timing is unknown. This
86 assumption predicts two scenarios in which a “rhythmic mode” can be restored (Zoefel &
87 VanRullen, 2017). First, if the timing of relevant events is known, the phase of the oscillation can
88 be adapted accordingly, and a loss of critical information during the low-excitability phase
89 avoided. As explained above, such an effect is fundamental for the field of “neural entrainment”
90 (Lakatos et al., 2008, 2019). A second scenario remained unexplored and was tested here: The
91 temporary suppression of input processing during the low-excitability phase can be tolerated if
92 expected events are irrelevant to perform a task, even if their timing is unpredictable. In this
93 scenario, the processing of irrelevant (but not relevant) events would be modulated by the
94 oscillatory phase.

95 We measured participants' EEG and asked them to detect pure tones at one sound
96 frequency (task-relevant tone) and ignore pure tones at another sound frequency (task-irrelevant
97 tone), presented at unpredictable moments (Figure 1A). We predicted that the processing of the
98 task-irrelevant, but not that of the task-relevant tone, depends on the phase of neural oscillations.
99 In a condition where both tones needed to be detected, we tested whether the presence of multiple
100 task-relevant tones leads to a rhythmic alternation of attentional focus between these tones – and
101 consequently, a phasic modulation of detection even for task-relevant tones – as previously
102 demonstrated for the visual system (Fiebelkorn et al., 2013; Helfrich et al., 2018).

103

104

105

Materials and Methods

106 Participants

107 Thirty native French participants took part in the experiment with informed consent for a monetary
108 reward of €25. The data of one participant was excluded due to technical issues, thus 29
109 participants (21 females, mean age = 22.34, SD = 1.2) were included in the final data analyses. All
110 experimental procedures were approved by the CPP (Comité de Protection des Personnes) Ouest
111 II Angers (protocol number 21.01.22.71950 / 2021-A00131-40).

112

113 Experimental Design

114 Participants performed a tone-in-noise detection task where they were presented with pure tones
115 at two different sound frequencies (440 Hz and 1026 Hz), embedded at unpredictable moments
116 into a continuous stream of pink noise (Figure 1A). They were instructed to press a button when
117 they hear a tone at the to-be-attended, task-relevant sound frequency and ignore the other one. A
118 correct detection was defined as a button press within 1 second after pure tone onset throughout
119 the experiment. All tones were 20ms in duration with a rise-and-fall period of 5ms. The continuous
120 pink noise was presented at ~70 dB SPL. Prior to the main experiment, the sound level of the pure
121 tones was titrated individually so that ~50% of tones were detected in the main task (see Adaptive
122 Staircase Procedure). In total, 504 pure tones at each sound frequency were presented. These were
123 divided into 12 experimental blocks, each ~ 5 min long.

124 In “selective attention” blocks, participants had to detect tones at one of the two sound
125 frequencies and to ignore the other. In “divided attention” blocks, they had to detect tones at both
126 sound frequencies. The order of the tones was pseudo-randomized with the constraint of a
127 transitional probability between 0.24 and 0.26. The stimulus-onset asynchrony between tones was
128 randomized between 2 and 5s with a uniform distribution to ensure temporal unpredictability. The
129 unpredictability of the tones prevented potential preparatory responses to the upcoming stimulus.
130 We adopted a rolling adaptive procedure to ensure that participants would detect the tone at
131 threshold level (50%) throughout the experiment. After each block, if the participant’s detection
132 probability was lower than 40% or higher than 60 %, the sound level of the tone at the
133 corresponding pitch was increased or decreased by 1dB, respectively. The block order (selective
134 attention – low pitch, selective attention – high pitch, divided attention) was counterbalanced
135 between participants.

136 Stimulus presentation was done via Matlab 2019a (MathWorks, Inc., Natick, USA) and
137 Psychtoolbox (Brainard, 1997). The auditory stimuli were presented using Etyomic ER-2 inserted
138 earphones and a Fireface UCX soundcard. The same sound card was used to send triggers to the
139 EEG system, ensuring synchronisation between sound and EEG.

140

141 Adaptive Staircase Procedure

142 Individual detection thresholds were determined separately for each of the two pure tones with a
143 1-up-1-down adaptive staircase procedure as implemented in the Palamedes toolbox (Prins &
144 Kingdom, 2018). In each adaptive trial, one pure tone was embedded randomly between 0.5 and
145 4.5s after the onset of a 5-second pink noise snippet. The participant had to press a button as soon
146 as they detected the pure tone. With a starting value of -30 dB of the total soundcard output (i.e.,
147 around 70 dB SPL), the sound level of the tone decreased in steps of 1 dB if the participant
148 correctly detected it, or increased accordingly if they missed the pure tone. The adaptive procedure
149 ended after 10 reversals, and the final 6 reversals were used to calculate the threshold. The
150 convergence of the staircase procedure was examined by visual inspection to determine whether
151 the threshold would be used in the following main experiment. If convergence failed, the adaptive
152 procedure was repeated. The average thresholds for high- and low frequency tones were -39.67 dB
153 (SD = 1.21 dB) and -37.10 dB (SD = 1.26 dB), respectively, resulting in ~ 50% detected tones
154 during both selective and divided attention (Figure 1B).

155

156 EEG Recording and Data Processing

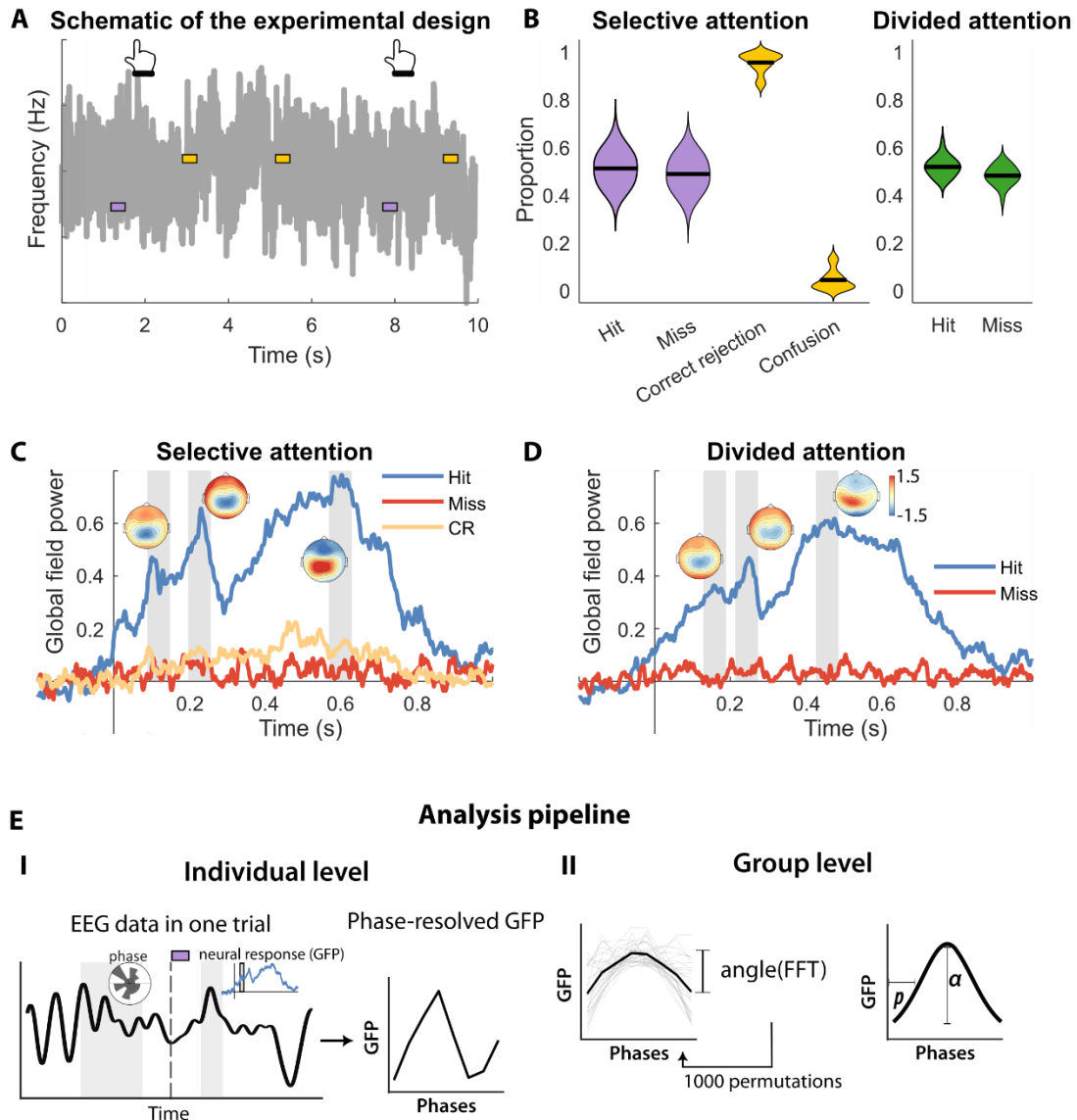
157 EEG was recorded using a Biosemi Active 2 amplifier (Biosemi, Amsterdam, Netherlands). 64
158 active electrodes positioned according to the international 10-10 system. The sampling rate of the
159 EEG recording was 2048 Hz. Equivalent to typical reference and ground electrodes, the Biosemi
160 system employs a “Common Mode Sense” active electrode and a “Driven Right Leg” passive
161 electrode located in the central-parietal region for common mode rejection purposes. The signal
162 offsets of all electrodes were kept under 50 μ V.

163 All EEG pre-processing steps were conducted using Matlab 2021a (MathWorks, Inc.,
164 Natick, USA) and the fieldtrip toolbox (Oostenveld et al., 2011). EEG data were re-referenced to
165 the average of all electrodes. Then, the data were high- and low-pass filtered (4th order Butterworth
166 filter, cut-off frequencies 0.5 Hz and 100 Hz, respectively). Noisy EEG channels were identified
167 by visually inspection and interpolated. Artefacts such as eye blinks, eye movements, muscle
168 movements, and channel noise were detected in an independent component analysis (ICA) applied
169 to down-sampled EEG data with a sampling rate of 256 Hz. Contaminated components were
170 detected by visual inspection and removed from data at the original sampling rate. The continuous
171 EEG data were segmented from -2s to +2s relative to each tone onset, termed “trials” in the
172 following. Trials with an absolute amplitude that exceeded 160 μ V were rejected.

173 We did not measure participants’ subjective perception of task-irrelevant tones as this
174 would have rendered them relevant. Instead, we used a neural proxy to infer how readily these
175 tones were processed, and how processing depended on pre-stimulus phase. In line with previous
176 work (Busch & VanRullen, 2010), we used global field power (GFP) evoked by tones as such a
177 proxy. For this purpose, event-related potentials (ERPs) were calculated for each participant,
178 separately for correctly detected (hits) and missed targets (misses) and for each condition in the 2
179 x 2 design (task-relevant vs task-irrelevant, selective vs divided attention). For the selective
180 attention condition, ERPs for trials where participants correctly did not respond to the task-

181 irrelevant tone (correct rejection; CR) were also calculated. GFP was extracted as the low-pass
182 filtered (cut-off frequency 10 Hz) standard deviation of the ERPs across EEG channels. Three
183 relevant time lags for tone processing were determined as local maxima (i.e. peaks identified with
184 the “findpeaks” Matlab function) in the grand average GFP from 0 to 1s after tone onset, separately
185 for selective and divided conditions (Figure 1C). As the aim of this step is the identification of
186 relevant time lags for tone processing, we restricted the analysis to detected task-relevant tones
187 (Figure 1C, D). Time windows of interest for the analysis of phasic effects (see below) were
188 selected as +/- 30ms around each of these three peaks. Single-trial GFP amplitudes were obtained
189 by averaging the GFP amplitude across time points within each time window of interest. This was
190 done separately for each experimental condition, including those without a behavioural response
191 (i.e. the task-irrelevant conditions).

192 We used a fast Fourier transform (FFT) with hanning tapers and sliding windows (0.02 s
193 steps) to extract EEG phases at frequencies from 2 Hz to 20 Hz (1 Hz steps) from single trials and
194 channels. The window size for phase estimation was linearly spaced between 2 (for 2 Hz) and 5.6
195 (for 20 Hz) cycles of the corresponding frequency. The subsequent analytical steps were restricted
196 to phases estimated from windows that do not include post-stimulus EEG data (cf. Figure 2A).
197 This avoid a potential contamination with stimulus-evoked responses that can lead to spurious
198 phase effects (Vinao-Carl et al., 2024; Zoefel & Heil, 2013).



199

200 *Figure 1.* A) Schematic of the tone-in-noise detection task. Purple and yellow rectangles denote task-relevant and
 201 task-irrelevant tones, respectively. In the main experiment, low and high tones served as task-relevant and task-
 202 task-irrelevant tones in different blocks. Grey line shows the continuous pink noise. B) Behavioural performance for
 203 selective attention (left) and divided attention (right) conditions. Black lines show the mean across participants. C,D)
 204 Global field power (GFP) for hit (blue; relevant tones), miss (red; relevant tones), and correct rejection (CR; yellow;
 205 irrelevant tones) in the selective (C) and divided (D) attention conditions. Grey areas indicate the time window selected
 206 for the phase dependence analysis. Insets show topographies of GFP at each time window for hit trials. E) Illustration
 207 of the analysis pipeline for the phase-dependence analysis. E-I) Extraction of single-trial phase estimates for individual
 208 participants. GFP in each phase bin was calculated to create the phase-resolved GFP values. E-II) The analysis
 209 procedure on the group level with simulated individual phase-resolved GFP for illustration (thin grey lines). The
 210 hypothesized phase effect was quantified by fitting a sine function to the averaged data (bold black line) and
 211 contrasting the amplitude a of this fit against that obtained in a permutation distribution ($N = 1000$) This analysis
 212 assumes that the phase p of individual sine functions is consistent across participants, an assumptions that we verified
 213 statistically (see Materials and Method; Results).

214 **Statistical analysis**

215 To address our main hypothesis, we tested whether the magnitude of the stimulus-evoked response
 216 (as GFP; see previous section) varies with pre-stimulus neural phase (Figure 1E). We used a
 217 statistical approach that a previous simulation study (Zoefel et al., 2019) showed to be particularly
 218 sensitive to such phasic effects (“sine fit binned” method in that study). For each condition,
 219 participant, EEG channel, frequency and time point separately, single trials were divided into 8
 220 equally spaced bins according to their phase (Figure 1E-I) and the average GFP amplitude
 221 extracted for each phase bin. We then fitted a sine function to the resulting phase-resolved GFP
 222 amplitude (Figure 1E-II). The amplitude of this sine function (a in Figure 1E-II) indexes how
 223 strongly tone processing is modulated by EEG phase whereas its phase (p in Figure 1E-II) reflects
 224 “preferred” and “non-preferred” phases for GFP (leading to highest and lowest GFP, respectively).
 225 To quantify phase effects statistically, we compared sine fit amplitudes with those obtained in a
 226 simulated null distribution, i.e. in the absence of a phasic modulation of tone processing. This null
 227 distribution was obtained by randomly assigning EEG phases to single trials and recomputing the
 228 amplitude of the sine 1000 times for each condition, EEG channel, frequency and time point
 229 (VanRullen, 2016a). For each combination of these factors, the sine amplitude from the original
 230 data was compared with the null distribution to obtain group-level z-scores:

$$231 \quad z = (a - \mu) / \sigma$$

232 where z reflects the group level effect in the original data, a is the amplitude value in the original
 233 data, and μ and σ are mean and standard deviation (across permutations) of the subject-averaged
 234 amplitude in the surrogate distribution, respectively. Z-scores were then converted to p-values
 235 (e.g., $z = 1.645$ would correspond to a significance threshold of $\alpha = 0.05$, one-tailed) and corrected
 236 for multiple comparisons using the false discovery rate (FDR). Finally, clusters in combinations
 237 of frequency, time, and EEG channel for the FDR-corrected p-values were identified using the
 238 “findcluster” function in the fieldtrip toolbox.

239 One advantage of the statistical method used is that it makes explicit assumptions on
 240 whether participants have consistent “preferred” EEG phases, reflected in the phase of the sine
 241 fitted to individual participants (Figure 1E-II). If these phases are uniformly distributed (i.e.
 242 inconsistent across participants), the sine fit amplitude is extracted separately for each participant
 243 and then averaged before the comparison with the surrogate distribution. In this way, the z-score
 244 defined above is independent of individual preferred EEG phases. If phases are non-uniformly
 245 distributed (i.e. consistent across participants), the phase-resolved GFP (Figure 1E-I) is first
 246 averaged across participants and the sine function is fitted to the resulting average (Figure 1E-II)
 247 before the comparison with the surrogate distribution. In this way, the z-score is only high when
 248 its phase is consistent across participants. To test which version of the test is appropriate in our
 249 case, we applied a Rayleigh’s test for circular uniformity (Circular Statistics Toolbox; Berens,
 250 2009) to the distribution of individual preferred EEG phases at each time-frequency point. We
 251 found a pre-stimulus cluster of significant phase consistency across participants (cf. Results), and
 252 adapted our statistical method accordingly (using the second version described).

253 We adapted this statistical approach to test whether task-relevant and task-irrelevant tones differ
254 in their phasic modulation. In this version, we contrasted the difference in averaged sine fit
255 amplitudes between the two conditions (relevant vs irrelevant) with another surrogate distribution
256 for which the condition label was randomly assigned to trials. This procedure yielded another z-
257 score which was calculated as described above.

258 In the divided attention condition, we additionally tested whether the processing of the low
259 and high frequency tone has a different preferred phase by comparing the phase difference between
260 the two tone conditions against zero (circular one-sample test against angle of 0; `circ_mtest.m` in
261 the Circular Statistics Toolbox).

262

263 **Source localisation of the phase-dependence effect**

264 We also explored the neural origins of the effects found in the analysis of EEG phase effects, using
265 standard procedures implemented in the fieldtrip toolbox. For this purpose, we used a standard
266 volume conduction model and electrode locations to calculate the leadfield matrix (10-mm
267 resolution). Then, for the selective attention condition, we calculated a spatial common filter using
268 the LCMV beamformer method ($\lambda = 5\%$; Van Veen et al., 1997) on the 20-Hz low-pass
269 filtered EEG data from -1s to -0.5s relative to tone onset. The chosen time window encompasses
270 all of the observed phase effects (cf. Results). This resulted in 2,015 source locations that were
271 inside the brain.

272 Single-trial EEG data from individual participants were projected onto the source space
273 with the spatial common filter. The analysis of phasic effects was then applied to data from each
274 source location as described above for the sensor level. Due to the large computational demand,
275 we used 100 permutations for the construction of surrogate distributions (z-score defined above),
276 a number shown to be sufficient in the past (VanRullen, 2016a). The voxels with the 1% largest z-
277 scores were selected as the origin of the corresponding effects on the sensory level. Note that, due
278 to the low spatial resolution of EEG, we explicitly treat these source-level results as explorative.

279

Results

280 Overview

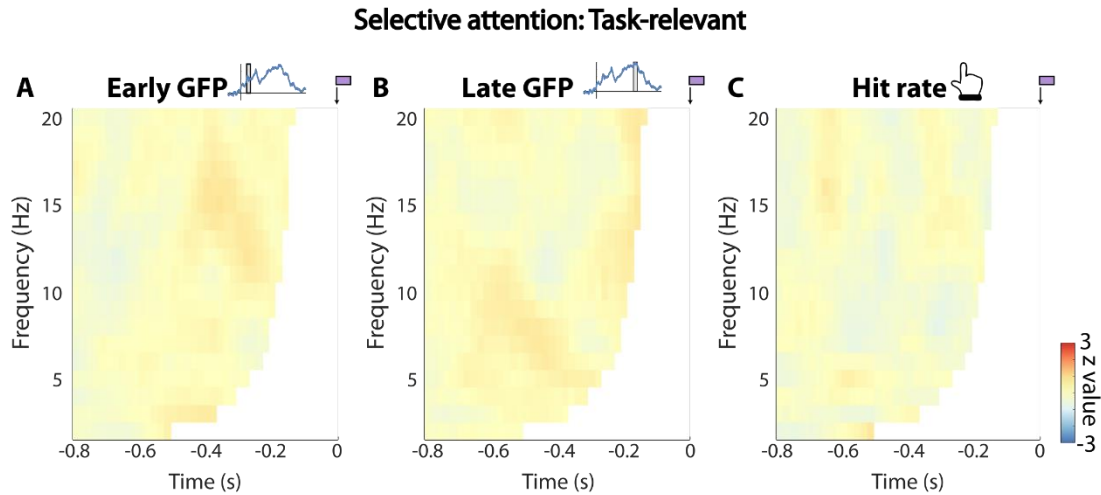
281 Participants were presented with tones at two different sound frequencies (Figure 1A). In some
282 experimental blocks, they were asked to detect one of them (task-relevant tone in the selective
283 attention condition) and ignore the other (task-irrelevant tone in the selective attention condition).
284 In other blocks, they were asked to detect both of them (divided attention condition).

285 On average, participants detected 51.18% (SD = 0.07%) and 51.84% (SD = 0.04%) of task-
286 relevant tones during selective and divided attention, respectively (Figure 1B), demonstrating
287 successful titration of individual thresholds (see Materials and Methods).

288 During both attentional conditions, task-relevant tones produced a strong increase in global
289 field power (GFP) if they were detected but not if they were missed (Figure 1C, D). We therefore
290 used the grand-average evoked GFP as a proxy for tone processing, and identified three time lags
291 with local GFP maxima for further analyses (grey in Figure 1C, D). The time lags for “early”,
292 “medium” and “late” evoked GFP were 119 ms, 227 ms and 598 ms for the selective attention
293 condition, and 159 ms, 243 ms, and 457 ms for the divided attention condition, respectively. We
294 used GFP as a principal measure of tone processing due to the lack of behavioural response to
295 task-irrelevant tones which would otherwise have rendered them relevant. Validating this measure
296 of neural processing, the GFP at each of the three time lags was significantly larger for detected
297 than for missed task-relevant tones during both selective (early: $t_{28} = 7.81$, $p < .001$; medium: t_{28}
298 $= 7.89$, $p < .001$; late: $t_{28} = 10.67$, $p < .001$) and divided attention (early: $t_{28} = 7.46$, $p < .001$;
299 medium: $t_{28} = 7.22$, $p < .001$; late: $t_{28} = 8.44$, $p < .001$).

300 Having identified critical time lags of tone processing, we extracted the GFP at each of the
301 three lags evoked by single tones (including task-irrelevant ones) and tested how strongly GFP
302 depends on pre-stimulus EEG phase in the different conditions (task-relevant vs irrelevant,
303 selective vs divided attention). Following previous work (Lui et al., 2023; Zoefel et al., 2019), we
304 fitted a sine function to GFP as a function of EEG phase (Figure 1E), and used the amplitude of
305 this fit (a in Figure 1E-II) as a measure of phasic modulation strength. Statistical reliability of the
306 phase effects was tested by comparison with a simulated null distribution (as z-score; see Material
307 and Methods).

308 In the following, we illustrate results separately for task-relevant (Figure 2) and task-irrelevant
309 tones (Figure 3) in the selective attention condition, respectively, as well as for the divided
310 attention condition (Figure 4). We only display results for early and late GFP, as no phasic
311 modulation was found for the medium time lag in any of the conditions.



312

313 *Figure 2. Results for task-relevant tones in the selective attention condition. The colour shows how strongly GFP*
 314 *(A,B) and hit rate (C) depends on EEG phase, expressed relative to a surrogate distribution, and averaged across*
 315 *channels. Time 0 corresponds to tone onset. In A and B, insets show relevant time lags for the analysis (early GFP:*
 316 *+119ms ; late GFP: +598ms. Time-frequency points “contaminated” by post-stimulus data (which is “smeared” into*
 317 *pre-stimulus phase estimates during spectral analysis) are masked.*

318 **Neural response evoked by task-irrelevant but not task-relevant tones depends on phase of** 319 **neural oscillations during selective attention**

320 We found that pre-stimulus EEG phase did not predict GFP evoked by task-relevant tones at any
 321 of the three time lags (all $p > 0.05$ after FDR correction; Figure 2A,B). Consistent with this result,
 322 the probability of detecting these tones was independent of pre-stimulus phase (all $p > 0.05$ after
 323 FDR correction; Figure 2C). In contrast, both early (Figure 3A-C) and late (Figure 3D-F) GFP
 324 evoked by task-irrelevant tones depended on pre-stimulus phase.

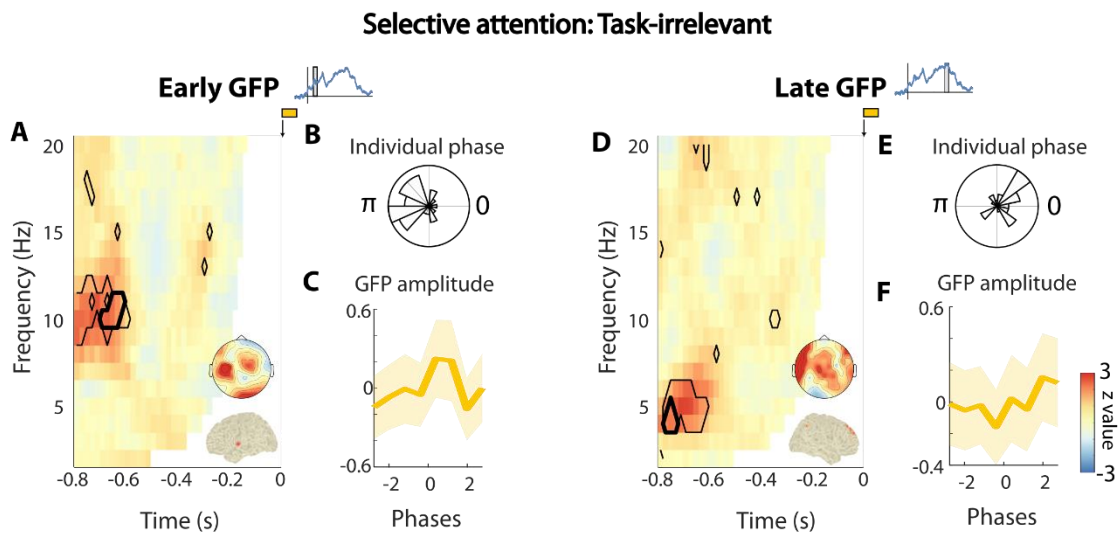
325 For the early lag, the phasic modulation was maximal at 10 Hz and 0.8 s preceding tone onset (z
 326 = 5.41, FDR-corrected $p = .003$). The EEG phase leading to maximal GFP at that time-frequency
 327 point was consistent across participants (Rayleigh’s test; $z = 6.21$, FDR corrected $p = .006$; Figure
 328 3B). The largest cluster of significant z-scores (FDR-corrected $p < 0.05$) was identified at ~10-11
 329 Hz, in the left central channels, and between -0.7s and -0.62s relative to tone onset (summed $z =$
 330 63.2, 14 time-frequency-channel points; Figure 3A). Explorative source localisation revealed that
 331 the phasic modulation originated from the left superior temporal cortex (Figure 3A, inset).

332 For the late lag, the phasic modulation was maximal at 5 Hz and 0.7 s preceding tone onset
 333 ($z = 5.49$, FDR-corrected $p = .001$). The EEG phase leading to maximal GFP at that time-frequency
 334 point was also consistent across participants ($z = 6.24$, FDR corrected $p = .006$; Figure 3E). The
 335 largest cluster of significant z-scores was identified at ~4-5 Hz and between -0.78 and -0.74s
 336 relative to tone onset (summed $z = 47.45$, 11 time-frequency-channel points; Figure 3D. This effect
 337 was localised to the right superior frontal gyrus and, to a lesser extent, the right inferior parietal
 338 cortex (Figure 3D, inset).

339

340 Contrasting amplitudes of the fitted sine functions between task-relevant and task-
 341 irrelevant tones, we found a stronger phasic modulation for the task-irrelevant tones at their
 342 relevant time-frequency points (Figure 3A,D) that concerned both early GFP ($z = 4.08, p < .001$;
 343 paired t-test) and late GFP ($z = 2.92, p = .004$). However, neither of these outcomes survived
 344 correction for multiple comparison ($p > 0.05$ after FDR correction).

345 Together, our results confirm previous findings that the processing of task-relevant
 346 auditory information is independent of the phase of neural oscillations (Zoefel & Heil, 2013), and
 347 extend them by demonstrating that such a phasic modulation reappears when the information is
 348 made irrelevant. Both alpha and theta oscillations, through their correspondence with different
 349 stages of neural processing, seem to contribute to rhythmic effects on unattended information
 350 during selective attention.



351
 352 *Figure 3. Results for task-irrelevant tones in the selective attention condition. A, D) Same as Figure 2A,B, but for*
 353 *task-irrelevant tones, and for channels selected for their significant phasic modulation of GFP ($p < .05$ after FDR*
 354 *correction). Black contours show the time-frequency points with significant phase effects. Bold black contours show*
 355 *the cluster with the largest summed z-score. Upper insets on the two panels show the topographical maps of z-scores*
 356 *in the corresponding time-frequency clusters. Lower insets show the 1% voxels with the largest source-projected z-*
 357 *scores in the same clusters. B, E) Distribution of individual phases of the sine function fitted to phase-resolved GFP*
 358 *(p in Figure 1E-II), at the time-frequency-channel combination with strongest phasic modulation (B: 11 Hz, -0.64s,*
 359 *C5; E: 4 Hz, -0.76s, FT7). C, F) GFP as a function of EEG phase from the same time-frequency-channel combination.*
 360 *The bold line shows the group-level average, the shaded area shows its standard error. Insets next to the titles show*
 361 *the GFP from Figure 1C with the time windows at which the analysis was performed.*

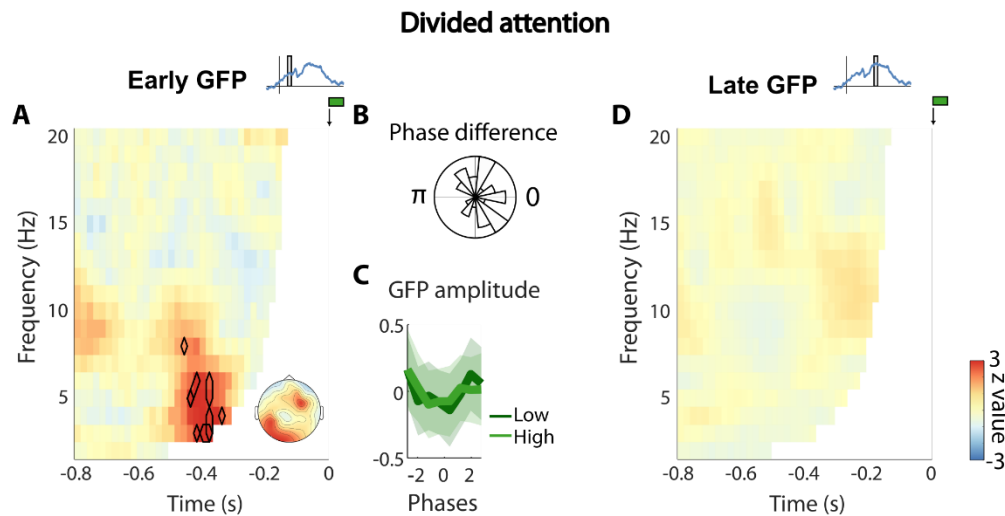
362
 363 **Early but not late response evoked by task-relevant tones depends on phase of neural**
 364 **oscillations during divided attention**

365 In the divided attention condition, only task-relevant tones were present. According to our
 366 principal hypothesis, the auditory system should suppress oscillations and instead operate in a
 367 continuous mode of processing to avoid a loss of information at the low-excitability phase.

368 However, an alternative possibility is that the presence of multiple target tones requires a rhythmic
 369 alternation of attentional focus between these tones as previously demonstrated for the visual
 370 system (Fiebelkorn et al., 2013; Helfrich et al., 2018). Such a case would lead to a phasic
 371 modulation of tone processing, similarly to what we observed for task-irrelevant tones in the
 372 selective attention condition.

373 Figure 4 shows how strongly the evoked GFP at early (A) and late (D) time lags depended on pre-
 374 stimulus EEG phase in the divided attention condition. We found a phasic modulation of tone
 375 processing only for the early time lag. This effect was maximal at 3 Hz and 0.42 s preceding tone
 376 onset ($z = 5.09$, FDR-corrected $p = .01$). However, we could not identify a cluster of significant z -
 377 scores, suggesting that these did not conglomerate in neighbouring electrodes, frequency, or time
 378 as evidently as for the selective attention condition. EEG phase leading to the strongest early GFP
 379 were similar for low- and high-frequency tones (Figure 4C), supported statistically by a
 380 distribution of their phase difference (Figure 4B) that did not significantly differ from zero (mean
 381 angle = 0.23, $p = .71$; circular one-sample test against angle of 0). The probability of detecting
 382 tones did not depend on pre-stimulus phase during divided attention (all FDR corrected $p < 0.05$;
 383 results for time-frequency point with strongest effect in Figure 4A: $z = 0.89$, $p = 0.37$).

384 Together, our results show that a rhythmic mode of processing reappears in the auditory system
 385 when confronted with multiple targets, but only affects early stages of target processing. In the
 386 presence of two target tones, the frequency of modulation is approximately divided by half as
 387 compared to a single tone, and the two target tones have similar preferred EEG phases for their
 388 processing. These results speak for a mechanism processing each of the two tones at successive
 389 cycles of a faster rhythm, as we explain in the Discussion.



390

391 *Figure 4. Results for task-relevant tones in the divided attention condition.* All conventions as in Figure 3, apart from
 392 panel B, which illustrates the distribution of phase difference between low- and high-frequency tones, and panel C,
 393 where results are shown separately for the low- and high frequency tones.

394

Discussion

395 The current study aimed to unveil the rhythm of auditory perception during selective and divided
396 attention. To this end, we asked participants to perform a target-in-noise detection task where they
397 had to attend to tones at one sound frequency and ignore another (selective attention), or had to
398 attend to both (divided attention).

399 In line with previous work (Zoefel & Heil, 2013; VanRullen et al., 2014) and our own
400 hypothesis, we found that neural and behavioural responses to task-relevant tones do not depend
401 on the pre-stimulus phase of neural oscillations during selective attention. Conversely, early and
402 late neural responses to task-irrelevant tones were modulated by the phase of pre-stimulus alpha
403 and theta oscillations, respectively. These results demonstrate that while neural oscillations seem
404 to be suppressed during attentive selection of single auditory targets, there exists a rhythmic mode
405 of perception in the auditory system that is applied to unattended sensory information. Finally, we
406 found evidence that this mode is also active when confronted with multiple auditory targets,
407 although restricted to early stages of their processing.

408

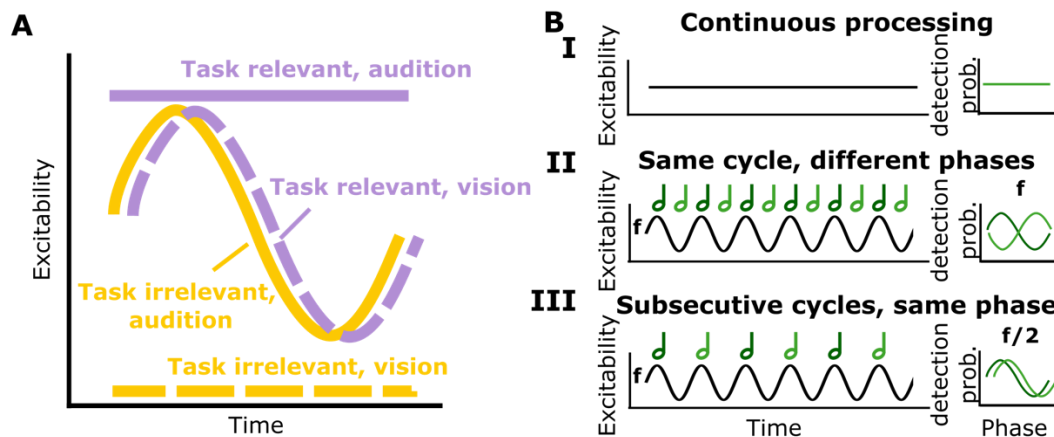
409 An inattentional rhythm in audition

410 It is a striking difference between modalities that selective attention *increases* the effect of neural
411 phase on the processing of temporarily unpredictable targets in the visual domain (Busch &
412 VanRullen, 2010) but *decreases* it in the auditory one (Zoefel & Heil, 2013; current study).
413 Confirming previous speculations (Zoefel & VanRullen, 2017), we here demonstrate that a
414 rhythmic mode of auditory processing is restored when stimuli become irrelevant and information
415 loss is tolerable. This “inattentional rhythm” that seems specific to audition may arise from specific
416 requirements on the auditory system during dynamic stimulus processing.

417 In contrast to the relatively stable visual environment, auditory inputs are often transient
418 and dynamic. Therefore, periodic sampling of the external environment may be more detrimental
419 for audition when temporarily unpredictable information is important for goal-directed behaviour.
420 In this case, the auditory system may engage in a desynchronised cortical state in the auditory
421 cortex that is associated with the active processing of incoming sensory inputs (Pachitariu et al.,
422 2015). As much as this “continuous mode” prevents the loss of information by suppressing
423 periodic moments of low excitability, it is likely to be metabolically demanding (Schroeder &
424 Lakatos, 2009). Therefore, the auditory system may limit the use of such a mode to scenarios in
425 which a loss of information is likely (such as the expectation of relevant events whose timing
426 cannot be predicted). This notion can also explain the prevalence of rhythm in acoustic information
427 (music, speech etc.): If relevant events are presented regularly, then their timing can be predicted
428 and the oscillatory phase adapted accordingly (Lakatos et al., 2008). Such a mechanism would
429 enable a “rhythmic mode” of processing even for task-relevant stimuli.

430 Based on these results, we propose that – due to its highly dynamic environment – the auditory
431 system always needs to be “one degree more attentive” to sensory information than the visual one.
432 We illustrate this idea in Figure 5A that can be summarised as follows: In the presence of

433 temporarily unpredictable, relevant information, the auditory system needs to operate in a
 434 continuous mode of high-excitability, whereas the visual system can sample rhythmically, due to
 435 the significantly slower input dynamics. A similar rhythmic mode of processing is used in the
 436 auditory system to sample unattended input, whereas it is processed in a mode of continuous low
 437 sensitivity in the visual case. The latter explains why we observed a phasic modulation of task-
 438 irrelevant tones in the current study, in contrast to an absence of such an effect in the visual
 439 modality (Busch & VanRullen, 2010). Our model is also supported by the finding that auditory
 440 distractors are more distracting than visual distractors (Berti & Schröger, 2001), even when the
 441 primary task is in the visual modality (Lui & Wöstmann, 2022). This might be because the auditory
 442 system exhibits periodic moments of high sensitivity to distractors and is therefore also more
 443 sensitive to potentially threatening stimuli that warrant immediate action.



444
 445 *Figure 5. Hypothetical “modes” of processing that do or do not rely on the phase of neural oscillations during*
 446 *selective (A) and divided attention (B). A) If the timing of relevant events is unknown, the auditory system might need*
 447 *to suppress neural oscillations to avoid a loss of information at the low-excitability phase, and operate in a mode of*
 448 *continuous high excitability (continuous purple line), whereas the visual system can operate rhythmically (dashed*
 449 *purple), due to its slower sensory dynamics. If events become irrelevant, the auditory system might change to a mode*
 450 *of periodic high sensitivity, reflected in a rhythmic sampling of irrelevant information (continuous yellow). The visual*
 451 *system might not need these high-sensitivity moments for irrelevant information, resulting in a continuous mode of*
 452 *low excitability (dashed yellow). B) Three hypothetical modes of processing during auditory divided attention. When*
 453 *multiple targets need to be processed, the auditory system might operate in a continuous mode of processing to avoid*
 454 *loss of information at a low-excitability phase (I, left). Such a mode would lead to a detection of these targets that is*
 455 *independent of phase (I, right). Alternatively, the presence of multiple targets might require an alternation of*
 456 *attentional focus between possible sound frequencies that relies on neural phase at the frequency f . This might be*
 457 *achieved by prioritizing different sound frequencies at different neural phases (II, left), leading to a target detection*
 458 *probability that depends on the phase at f , and a preferred phase for detection that changes with sound frequency of*
 459 *the target (II, right). In an alternative rhythmic mode, possible sound frequencies are processed at the same (high-*
 460 *excitability) phase of f , but in subsecutive cycles (III, left), leading to a phase effect at $f/2$ and to similar preferred*
 461 *phases across sound frequencies (III, right). The latter is what we have observed in the current study (cf. Figure 4A).*

462
 463 **Alpha and theta oscillations modulate distinct processing steps of irrelevant events**

464 We found that the pre-stimulus phase of alpha oscillations predicts a relatively early response to
 465 task-irrelevant tones whereas the pre-stimulus phase of theta oscillations predicts later responses

466 (Figure 3). We speculate that this finding can be attributed to distinct steps in the processing of
467 task-irrelevant events that depend on different oscillatory frequency bands.

468 The phase of alpha oscillations is posited to gate perception via pulsed inhibition (Jensen
469 & Mazaheri, 2010) at early stages of cortical processing where the encoding of sensory events
470 takes place (Klimesch et al., 2011). Indeed, the phasic modulation of the early evoked response in
471 the alpha band seemed to originate from relatively early stages of the auditory cortical hierarchy
472 in our study (Figure 3B). The timing of the early evoked GFP (~119 ms) is well in line with
473 components of stimulus-evoked neural responses (e.g., P1, N1) that have been associated with
474 stimulus encoding (Näätänen & Picton, 1987). Although imaging methods with higher spatial
475 localisation are required to validate this hypothesis, we speculate that alpha oscillations phasically
476 modulate the encoding of task-irrelevant events (e.g., distractors).

477 Stimulus-evoked neural responses at later delays have been associated with higher-level
478 cognitive operations, such as distractibility (Chao & Knight, 1995) as well as response execution
479 and inhibition (Bokura et al., 2001). Theta oscillations in the frontal cortex have been considered
480 a neural proxy of executive control (Mizuhara & Yamaguchi, 2007; Sauseng et al., 2007). A
481 previous study showed evidence for a theta rhythm in distractibility by showing that perceptual
482 sensitivity is explained by pre-distractor theta phase (Lui et al., 2023). It is thus possible that the
483 propensity to ignore task-irrelevant events depends on pre-stimulus theta oscillations. The later
484 timing of the theta-phase modulation in our study as well as its localisation to more frontal brain
485 regions is in line with this assumption (Figure 3D). This effect may therefore reflect the inhibition
486 of the processing of task-irrelevant events that occurs after their encoding. The fact that only early
487 phasic effects were present, but the later theta-phase modulation was absent during divided
488 attention (Figure 4) further supports this assumption, as no distractors needed to be inhibited in
489 that condition.

490 It remains an open question why the strongest phase effect occurred relatively early before
491 tone onset (~-800 to -600 ms), and earlier than what has previously been reported (Busch &
492 VanRullen, 2010; Harris et al., 2018; Zazio et al., 2021). On the one hand, the closer to stimulus
493 onset, the stronger is the “contamination” of phase estimates by post-stimulus data (Vinao-Carl et
494 al., 2024; Zoefel & Heil, 2013), potentially obscuring maxima closer to tone onset. On the other
495 hand, the earliest time points that remain unaffected by temporal smearing can be estimated
496 precisely and do not show the strongest effects (Figures 2 – 4). Other factors might therefore play
497 a role and need to be identified in future work. For example, it is possible that the perception and
498 suppression of task-irrelevant auditory events is achieved through connectivity with other brain
499 regions that eventually cascades down to the auditory system at stimulus onset.

500

501 **A rhythmic mode in auditory divided attention**

502 We found evidence for a rhythmic mode of processing during auditory divided attention, and our
503 results provide insights into a mechanistic implementation of such a mode. The phasic modulation
504 of the early GFP evoked by the two tones (Figure 4A) contradicts our initial hypothesis that neural
505 oscillations are suppressed during divided attention to task-relevant tones (Figure 5B-I).

506 Nevertheless, the two tones (low- and high sound frequency) could be processed in the same
507 oscillatory cycle but at different phases (Figure 5B-II) as often proposed in the context of neural
508 oscillations (Gips et al., 2016; Jensen et al., 2014), or in subsecutive cycles (Gaillard & Ben
509 Hamed, 2022) and at a similar phase (Figure 5B-III). Based on predicted results patterns that can
510 distinguish these alternatives (Figure 5B, right panels), our results favour the second one, as (1)
511 the frequency of the early modulation is divided approximately by two as compared to the
512 processing of a single tone (compare Figure 3A and 4A); and (2) phases do not differ between the
513 low and high frequency tones (Figure 4B,C). Therefore, our results suggest that alpha oscillations
514 do not only modulate the processing of task-irrelevant information, but also the early stages of
515 task-relevant processing during divided attention, alternating between possible sound frequencies
516 of targets.

517 This conclusion is well in line with previous research. For instance, the frequency of visual
518 perception decreases with increasing number of to-be-attended features (Holcombe & Chen, 2013;
519 Schmid et al., 2022). The spotlight of attention has been posited to alternate between two locations
520 when both are attended, dividing an overall ~8 Hz rhythm into a ~4 Hz fluctuation in perceptual
521 sensitivity per location (Landau & Fries, 2012; Song et al., 2014; Zoefel & Sokoliuk, 2014). In the
522 auditory modality, a similar alternation between the two ears has been reported during divided
523 attention (Ho et al., 2017). We here extend this mechanism to an alternation between sound
524 frequencies, supporting the previous observation that oscillatory mechanisms follow the tonotopic
525 organisation of the auditory cortex (Lakatos et al., 2013; L’Hermite & Zoefel, 2023).

526

527 **Conclusion**

528 By showing that the processing of task-irrelevant but not task-relevant tones depends on
529 the pre-stimulus phase of neural oscillations during selective attention, we here provide evidence
530 that oscillatory mechanisms in audition critically depend on the degree of possible information
531 loss. We propose that this effect represents a crucial difference to the visual modality which might
532 not be equally responsive to sensory information (Figure 5). During divided attention, cycles of
533 alpha oscillations seem to alternate between possible targets similar to what was observed in
534 vision, suggesting an attentional process that generalises across modalities.

535

536

References

- 537 Berens, P. (2009). CircStat: A MATLAB Toolbox for Circular Statistics. *Journal of Statistical Software*,
538 31, 1–21. <https://doi.org/10.18637/jss.v031.i10>
- 539 Berti, S., & Schröger, E. (2001). A comparison of auditory and visual distraction effects: Behavioral and
540 event-related indices. *Cognitive Brain Research*, 10(3), 265–273. [https://doi.org/10.1016/S0926-6410\(00\)00044-6](https://doi.org/10.1016/S0926-6410(00)00044-6)
541
- 542 Bokura, H., Yamaguchi, S., & Kobayashi, S. (2001). Electrophysiological correlates for response
543 inhibition in a Go/NoGo task. *Clinical Neurophysiology*, 112(12), 2224–2232.
544 [https://doi.org/10.1016/S1388-2457\(01\)00691-5](https://doi.org/10.1016/S1388-2457(01)00691-5)
- 545 Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, 10(4), 433–436.
546 <https://doi.org/10.1163/156856897X00357>
- 547 Busch, N. A., Dubois, J., & VanRullen, R. (2009). The Phase of Ongoing EEG Oscillations Predicts
548 Visual Perception. *Journal of Neuroscience*, 29(24), 7869–7876.
549 <https://doi.org/10.1523/JNEUROSCI.0113-09.2009>
- 550 Busch, N. A., & VanRullen, R. (2010). Spontaneous EEG oscillations reveal periodic sampling of visual
551 attention. *Proceedings of the National Academy of Sciences*, 107(37), 16048–16053.
552 <https://doi.org/10.1073/pnas.1004801107>
- 553 Chao, L. L., & Knight, R. T. (1995). Human prefrontal lesions increase distractibility to irrelevant sensory
554 inputs. *Neuroreport: An International Journal for the Rapid Communication of Research in*
555 *Neuroscience*, 6(12), 1605–1610. <https://doi.org/10.1097/00001756-199508000-00005>
- 556 Dugué, L., Marque, P., & VanRullen, R. (2011). The Phase of Ongoing Oscillations Mediates the Causal
557 Relation between Brain Excitation and Visual Perception. *The Journal of Neuroscience*, 31(33), 11889–
558 11893. <https://doi.org/10.1523/JNEUROSCI.1161-11.2011>
- 559 Dugué, L., Marque, P., & VanRullen, R. (2015). Theta Oscillations Modulate Attentional Search
560 Performance Periodically. *Journal of Cognitive Neuroscience*, 27(5), 945–958.
561 https://doi.org/10.1162/jocn_a_00755
- 562 Fiebelkorn, I. C., Saalman, Y. B., & Kastner, S. (2013). Rhythmic sampling within and between objects
563 despite sustained attention at a cued location. *Current Biology*, 23(24), 2553–2558.
564 <https://doi.org/10.1016/j.cub.2013.10.063>
- 565 Gaillard, C., & Ben Hamed, S. (2022). The neural bases of spatial attention and perceptual rhythms.
566 *European Journal of Neuroscience*, 55(11–12), 3209–3223. <https://doi.org/10.1111/ejn.15044>
- 567 Gips, B., van der Eerden, J. P. J. M., & Jensen, O. (2016). A biologically plausible mechanism for
568 neuronal coding organized by the phase of alpha oscillations. *European Journal of Neuroscience*, 44(4),
569 2147–2161. <https://doi.org/10.1111/ejn.13318>
- 570 Harris, A. M., Dux, P. E., & Mattingley, J. B. (2018). Detecting unattended stimuli depends on the phase
571 of prestimulus neural oscillations. *The Journal of Neuroscience*, 38(12), 3092–3101.
572 <https://doi.org/10.1523/JNEUROSCI.3006-17.2018>

- 573 Helfrich, R. F., Fiebelkorn, I. C., Szczepanski, S. M., Lin, J. J., Parvizi, J., Knight, R. T., & Kastner, S.
 574 (2018). Neural mechanisms of sustained attention are rhythmic. *Neuron*, *99*(4), 854-865.e5.
 575 <https://doi.org/10.1016/j.neuron.2018.07.032>
- 576 Henry, M. J., & Obleser, J. (2012). Frequency modulation entrains slow neural oscillations and optimizes
 577 human listening behavior. *Proceedings of the National Academy of Sciences*, *109*(49), 20095–20100.
- 578 Ho, H. T., Leung, J., Burr, D. C., Alais, D., & Morrone, M. C. (2017). Auditory sensitivity and decision
 579 criteria oscillate at different frequencies separately for the two ears. *Current Biology*, *27*(23), 3643-
 580 3649.e3. <https://doi.org/10.1016/j.cub.2017.10.017>
- 581 Holcombe, A. O., & Chen, W.-Y. (2013). Splitting attention reduces temporal resolution from 7 Hz for
 582 tracking one object to <3 Hz when tracking three. *Journal of Vision*, *13*(1), 12.
 583 <https://doi.org/10.1167/13.1.12>
- 584 Jensen, O., Gips, B., Bergmann, T. O., & Bonnefond, M. (2014). Temporal coding organized by coupled
 585 alpha and gamma oscillations prioritize visual processing. *Trends in Neurosciences*, *37*(7), 357–369.
 586 <https://doi.org/10.1016/j.tins.2014.04.001>
- 587 Jensen, O., & Mazaheri, A. (2010). Shaping Functional Architecture by Oscillatory Alpha Activity:
 588 Gating by Inhibition. *Frontiers in Human Neuroscience*, *4*. <https://doi.org/10.3389/fnhum.2010.00186>
- 589 Klimesch, W., Fellinger, R., & Freunberger, R. (2011). Alpha Oscillations and Early Stages of Visual
 590 Encoding. *Frontiers in Psychology*, *2*. <https://doi.org/10.3389/fpsyg.2011.00118>
- 591 Kubovy, M. (1988). Should we resist the seductiveness of the space: Time: Vision: Audition analogy?
 592 *Journal of Experimental Psychology: Human Perception and Performance*, *14*(2), 318–320.
 593 <https://doi.org/10.1037/0096-1523.14.2.318>
- 594 Lakatos, P., Gross, J., & Thut, G. (2019). A new unifying account of the roles of neuronal entrainment.
 595 *Current Biology*, *29*(18), R890–R905. <https://doi.org/10.1016/j.cub.2019.07.075>
- 596 Lakatos, P., Karmos, G., Mehta, A. D., Ulbert, I., & Schroeder, C. E. (2008). Entrainment of neuronal
 597 oscillations as a mechanism of attentional selection. *Science*, *320*(5872), 110–113.
 598 <https://doi.org/10.1126/science.1154735>
- 599 Lakatos, P., Musacchia, G., O’Connell, M. N., Falchier, A. Y., Javitt, D. C., & Schroeder, C. E. (2013).
 600 The spectrotemporal filter mechanism of auditory selective attention. *Neuron*, *77*(4), 750–761.
 601 <https://doi.org/10.1016/j.neuron.2012.11.034>
- 602 Landau, A. N., & Fries, P. (2012). Attention samples stimuli rhythmically. *Current Biology*, *22*(11),
 603 1000–1004. <https://doi.org/10.1016/j.cub.2012.03.054>
- 604 L’Hermite, S., & Zoefel, B. (2023). Rhythmic Entrainment Echoes in Auditory Perception. *Journal of*
 605 *Neuroscience*, *43*(39), 6667–6678. <https://doi.org/10.1523/JNEUROSCI.0051-23.2023>
- 606 Lui, T. K.-Y., Obleser, J., & Wöstmann, M. (2023). Slow neural oscillations explain temporal fluctuations
 607 in distractibility. *Progress in Neurobiology*, *226*, 102458.
 608 <https://doi.org/10.1016/j.pneurobio.2023.102458>
- 609 Lui, T. K.-Y., & Wöstmann, M. (2022). Effects of temporally regular versus irregular distractors on goal-
 610 directed cognition and behavior. *Scientific Reports*, *12*, 10020. <https://doi.org/10.1038/s41598-022-13211-3>
 611 13211-3

- 612 Mathewson, K. E., Gratton, G., Fabiani, M., Beck, D. M., & Ro, T. (2009). To See or Not to See:
 613 Prestimulus Phase Predicts Visual Awareness. *Journal of Neuroscience*, *29*(9), 2725–2732.
 614 <https://doi.org/10.1523/JNEUROSCI.3963-08.2009>
- 615 Mizuhara, H., & Yamaguchi, Y. (2007). Human cortical circuits for central executive function emerge by
 616 theta phase synchronization. *NeuroImage*, *36*(1), 232–244.
 617 <https://doi.org/10.1016/j.neuroimage.2007.02.026>
- 618 Näätänen, R., & Picton, T. (1987). The N1 Wave of the Human Electric and Magnetic Response to
 619 Sound: A Review and an Analysis of the Component Structure. *Psychophysiology*, *24*(4), 375–425.
 620 <https://doi.org/10.1111/j.1469-8986.1987.tb00311.x>
- 621 Obleser, J., & Kayser, C. (2019). Neural entrainment and attentional selection in the listening brain.
 622 *Trends in Cognitive Sciences*, *23*(11), 913–926. <https://doi.org/10.1016/j.tics.2019.08.004>
- 623 Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J.-M. (2011). FieldTrip: Open source software for
 624 advanced analysis of MEG, EEG, and invasive electrophysiological data. *Computational Intelligence and*
 625 *Neuroscience*, *2011*, 156869. <https://doi.org/10.1155/2011/156869>
- 626 Pachitariu, M., Lyamzin, D. R., Sahani, M., & Lesica, N. A. (2015). State-Dependent Population Coding
 627 in Primary Auditory Cortex. *The Journal of Neuroscience*, *35*(5), 2058–2073.
 628 <https://doi.org/10.1523/JNEUROSCI.3318-14.2015>
- 629 Prins, N., & Kingdom, F. A. A. (2018). Applying the model-comparison approach to test specific research
 630 hypotheses in psychophysical research using the palamedes toolbox. *Frontiers in Psychology*, *9*, 1250.
 631 <https://doi.org/10.3389/fpsyg.2018.01250>
- 632 Sauseng, P., Hoppe, J., Klimesch, W., Gerloff, C., & Hummel, F. C. (2007). Dissociation of sustained
 633 attention from central executive functions: Local activity and interregional connectivity in the theta range.
 634 *European Journal of Neuroscience*, *25*(2), 587–593. <https://doi.org/10.1111/j.1460-9568.2006.05286.x>
- 635 Schmid, R. R., Pomper, U., & Ansorge, U. (2022). Cyclic reactivation of distinct feature dimensions in
 636 human visual working memory. *Acta Psychologica*, *226*, 103561.
 637 <https://doi.org/10.1016/j.actpsy.2022.103561>
- 638 Schroeder, C. E., & Lakatos, P. (2009). Low-frequency neuronal oscillations as instruments of sensory
 639 selection. *Trends in Neurosciences*, *32*(1), 9–18. <https://doi.org/10.1016/j.tins.2008.09.012>
- 640 Song, K., Meng, M., Chen, L., Zhou, K., & Luo, H. (2014). Behavioral Oscillations in Attention:
 641 Rhythmic α Pulses Mediated through θ Band. *Journal of Neuroscience*, *34*(14), 4837–4844.
 642 <https://doi.org/10.1523/JNEUROSCI.4856-13.2014>
- 643 van Bree, S., Sohoglu, E., Davis, M. H., & Zoefel, B. (2021). Sustained neural rhythms reveal
 644 endogenous oscillations supporting speech perception. *PLOS Biology*, *19*(2), e3001142.
 645 <https://doi.org/10.1371/journal.pbio.3001142>
- 646 Van Veen, B. D., Van Drongelen, W., Yuchtman, M., & Suzuki, A. (1997). Localization of brain
 647 electrical activity via linearly constrained minimum variance spatial filtering. *IEEE Transactions on*
 648 *Biomedical Engineering*, *44*(9), 867–880. <https://doi.org/10.1109/10.623056>
- 649 VanRullen, R. (2016a). How to Evaluate Phase Differences between Trial Groups in Ongoing
 650 Electrophysiological Signals. *Frontiers in Neuroscience*, *10*. <https://doi.org/10.3389/fnins.2016.00426>

THE INATTENTIONAL RHYTHM IN AUDITION

- 651 VanRullen, R. (2016b). Perceptual Cycles. *Trends in Cognitive Sciences*, 20(10), 723–735.
652 <https://doi.org/10.1016/j.tics.2016.07.006>
- 653 VanRullen, R., Zoefel, B., & Ilhan, B. (2014). On the cyclic nature of perception in vision versus
654 audition. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 369(1641), 20130214.
655 <https://doi.org/10.1098/rstb.2013.0214>
- 656 Vinao-Carl, M., Gal-Shohet, Y., Rhodes, E., Li, J., Hampshire, A., Sharp, D., & Grossman, N. (2024).
657 Just a phase? Causal probing reveals spurious phasic dependence of sustained attention. *NeuroImage*,
658 285, 120477. <https://doi.org/10.1016/j.neuroimage.2023.120477>
- 659 Wöstmann, M., Lui, T. K.-Y., Frieze, K. H., Kreitewolf, J., Naujokat, M., & Obleser, J. (2020). The
660 vulnerability of working memory to distraction is rhythmic. *Neuropsychologia*, 146, 107505.
661 <https://doi.org/10.1016/j.neuropsychologia.2020.107505>
- 662 Zazio, A., Ruhnau, P., Weisz, N., & Wutz, A. (2021). Pre-stimulus alpha-band power and phase
663 fluctuations originate from different neural sources and exert distinct impact on stimulus-evoked
664 responses. *European Journal of Neuroscience*, ejn.15138. <https://doi.org/10.1111/ejn.15138>
- 665 Zoefel, B., Davis, M. H., Valente, G., & Riecke, L. (2019). How to test for phasic modulation of neural
666 and behavioural responses. *NeuroImage*, 202, 116175. <https://doi.org/10.1016/j.neuroimage.2019.116175>
- 667 Zoefel, B., & Heil, P. (2013). Detection of near-threshold sounds is independent of EEG phase in
668 common frequency bands. *Frontiers in Psychology*, 4. <https://doi.org/10.3389/fpsyg.2013.00262>
- 669 Zoefel, B., & Sokoliuk, R. (2014). Investigating the Rhythm of Attention on a Fine-Grained Scale:
670 Evidence from Reaction Times. *Journal of Neuroscience*, 34(38), 12619–12621.
671 <https://doi.org/10.1523/JNEUROSCI.2134-14.2014>
- 672 Zoefel, B., & VanRullen, R. (2017). Oscillatory mechanisms of stimulus processing and selection in the
673 visual and auditory systems: State-of-the-art, speculations and suggestions. *Frontiers in Neuroscience*, 11,
674 13.
- 675