

An oscillatory pipelining mechanism supporting reviewing during visual exploration and reading

Jensen, Ole; Pan, Yali; Frisson, Steven; Wang, Lin

DOI:

[10.1016/j.tics.2021.08.008](https://doi.org/10.1016/j.tics.2021.08.008)

License:

Creative Commons: Attribution-NonCommercial-NoDerivs (CC BY-NC-ND)

Document Version

Peer reviewed version

Citation for published version (Harvard):

Jensen, O, Pan, Y, Frisson, S & Wang, L 2021, 'An oscillatory pipelining mechanism supporting reviewing during visual exploration and reading', *Trends in Cognitive Sciences*, vol. 25, no. 12, pp. 1033-1044.
<https://doi.org/10.1016/j.tics.2021.08.008>

[Link to publication on Research at Birmingham portal](#)

Publisher Rights Statement:

This is the accepted author manuscript of the following article published in Trends in Cognitive Sciences:
<https://doi.org/10.1016/j.tics.2021.08.008>

General rights

Unless a licence is specified above, all rights (including copyright and moral rights) in this document are retained by the authors and/or the copyright holders. The express permission of the copyright holder must be obtained for any use of this material other than for purposes permitted by law.

- Users may freely distribute the URL that is used to identify this publication.
- Users may download and/or print one copy of the publication from the University of Birmingham research portal for the purpose of private study or non-commercial research.
- User may use extracts from the document in line with the concept of 'fair dealing' under the Copyright, Designs and Patents Act 1988 (?)
- Users may not further distribute the material nor use it for the purposes of commercial gain.

Where a licence is displayed above, please note the terms and conditions of the licence govern your use of this document.

When citing, please reference the published version.

Take down policy

While the University of Birmingham exercises care and attention in making items available there are rare occasions when an item has been uploaded in error or has been deemed to be commercially or otherwise sensitive.

If you believe that this is the case for this document, please contact UBIRA@lists.bham.ac.uk providing details and we will remove access to the work immediately and investigate.

[Click here to view linked References](#)

1 **An oscillatory pipelining mechanism supporting previewing during visual exploration**
2 **and reading**

3 Ole Jensen^{1*}, Yali Pan¹, Steven Frisson¹ and Lin Wang^{2,3}

4 ¹Centre for Human Brain Health, School of Psychology, University of Birmingham, U.K.

5 ²Department of Psychiatry and the Athinoula A. Martinos Center for Biomedical Imaging,
6 Massachusetts General Hospital, Harvard Medical School, Charlestown, Massachusetts
7 02129, US

8 ³Department of Psychology, Tufts University, Medford, Massachusetts 02155, US

9 *Corresponding author

10 **Keywords**

11 saccades, alpha oscillations, phase coding, preview, visual exploration, reading

12
13 **Abstract**

14 Humans have a remarkable ability to efficiently explore visual scenes and text employing eye
15 movements. Humans typically make eye movements (saccades) every ~250ms. Since saccade
16 initiation and execution take 100ms, this leaves only ~150ms to recognize the fixated object
17 (or word), while simultaneously previewing candidates for the next saccade goal. We propose
18 a *pipelining mechanism* where serial processing occurs within a specific brain region, whereas
19 parallel processing occurs across different brain regions. The mechanism is timed by alpha
20 oscillations that coordinate the saccades, visual recognition and previewing in the cortical
21 hierarchy. Consequently, the neuronal mechanism supporting natural vision and saccades must
22 be studied in unison to uncover the brain mechanisms supporting visual exploration and
23 reading.

24

25

26

27 Main text

28 The remarkable efficiency of visual exploration and reading

29 Our understanding of natural vision presents an intriguing conundrum: How do we manage to
30 efficiently explore visual scenes and text by eye movements given the relatively slow and
31 spatially limited processing capabilities of the human visual system? We saccade every 250 –
32 300 ms when reading and visually exploring natural scenes. Given that it takes about 100 ms
33 to initiate and execute a saccadic motor program [1, 2], there is only 150 – 200 ms available to
34 process the fixated object or word while in parallel planning the next saccade. Since saccades
35 typically land on informative objects or words [3, 4] (Figure 1), a *parafoveal previewing*
36 process is required when exploring and deciding on the next saccade goal. The deployment of
37 pre-saccadic attention has been the topic of many studies [5, 6]. In the context of saccadic
38 exploration, we need to uncover how the visual system can achieve this remarkable
39 computational feat. The fast computation must rely on a highly tuned machinery in which eye
40 movements are coordinated with the visual input [7]. We here suggest a new framework in
41 which visual exploration and reading are supported by a **pipelining** mechanism. This pipelining
42 mechanism is coordinated by neuronal oscillations that serve to organize the visual
43 representations in a temporal code and guide the information through the visual hierarchy.

44 The temporal constraints during visual exploration and reading

45 The recognition of currently fixated objects as well as for deciding on future saccade goals is
46 typically done within 150 ms after fixation onset. This is because saccades are initiated as often
47 as every 250 ms and it takes about 100ms to initiate and execute a saccadic program towards
48 the target [1, 2]. Moreover, it takes about 60 ms for information to travel from the retina to the
49 visual cortex [8], leaving about 90 ms for neocortical processing of the fixated object (Figure
50 1). Even with such high temporal constraints, it has been shown that it is possible to identify
51 visual objects at the semantic level (“meaning”, such as animacy features) within 150 ms.
52 Multivariate approaches applied to MEG data allow for identifying the *neuronal fingerprint*
53 associated with semantic features [9]. It was found that naturalness and animacy can be
54 decoded from multivariate brain activity at respectively 122 ms and 157 ms after stimulus onset
55 [9]. This timing is consistent with intracranial recordings in monkeys in which object category
56 was decoded within 125 ms in the inferior temporal cortex [10]. Therefore, existing studies
57 suggest that it is possible to identify visual objects at the semantic level (“meaning”) within
58 150 ms. Because future saccade goals also must be explored and selected within the same time
59 window, this poses a serious computational challenge to the visual system (Figure 1B). Due to

60 the head-start of processing from the parafoveal previewing prior to the saccade, recognition
61 of the fixated object is reduced to about 110 ms [1, 2, 11]. This then leaves ~40 ms more for
62 previewing the upcoming saccade goal (Figure 1C); i.e. it buys time to alter the saccade plan
63 if, for instance, the saccade goal is deemed uninformative. Therefore, the acceleration of visual
64 processing by previewing is likely to be essential for efficient visual exploration.

65
66 The temporal constraints during reading are equally tight. After the retinal input has arrived in
67 occipital cortex at 60 ms, the visual word form area (VWFA) is engaged in orthographic
68 processing at about 90 ms [12] (Figure 1D-E). Then follows lexical processing and semantic
69 recognition supported by the left inferior temporal cortex [12]. The lexical processing has been
70 shown to take up to 150ms. Eye-tracking research has demonstrated that fixation times are
71 longer for low- compared to high-frequency words [11, 13] and that the word-frequency effect
72 is present at least within 150 ms as revealed by survival analysis [14, 15] (Figure 1E).
73 Electrophysiological findings also support the notion that lexical identification can take up to
74 150 ms [16-19]. Since saccade initiation and execution take about 100 ms, there would be
75 almost no time to evaluate the next saccade goal [1, 20] (Figure 1E). Again, previewing will
76 serve to reduce the lexico-semantic recognition possibly to about 110 ms [1, 2, 11], which will
77 leave time to evaluate the next saccade goal (Figure 1F). This allows just enough time to alter
78 the saccade plan if the parafoveal word has already been processed to a sufficient degree that
79 a fixation is no longer necessary.

80

81 **How deeply are previewed objects processed?**

82 It is interesting to consider previewing in the context of parafoveal visual acuity. While acuity
83 drops and crowding effects become more prevalent for parafoveal vision (2–5 degrees relative
84 to the current fixation) [21], our eyes still saccade to relevant (and not necessarily salient) parts
85 of visual scenes [3, 4, 22-24]. Using gaze-contingent paradigms occluding the peripheral view,
86 it has been demonstrated that the effective visual span guiding saccades is about 8 degrees [25].
87 How deeply are previewed objects processed before we saccade to them? In support of
88 previewing at the semantic level, one study demonstrated that search times are faster for objects
89 that are inconsistent with their scene contexts (e.g. *a tube of toothpaste in the living room*) as
90 compared to scene-consistent objects (e.g. *a tube of toothpaste in the bathroom*) [26]. A recent
91 EEG study investigated fixation-related potentials (FRPs) in response to the fixation on pre-
92 target objects prior to saccading to target objects that were either consistent or inconsistent with
93 the context of the scenes. A larger negative potential at ~300 ms (akin to the N400 type ERP

94 effect) was observed in response to the pre-target when the target object (e.g. *a tube of*
95 *toothpaste*) was inconsistent with the scene (e.g. *living room*). These findings provide support
96 in favour of parafoveal processing at the semantic level [27].

97
98 The perceptual span during reading has also been investigated using gaze-contingent paradigms
99 in which text is occluded to the left and/or right of the gaze. These studies have shown that the
100 visual span extends 14–15 letter spaces (2–3 words) to the right of fixation and 3–4 letters to
101 the left [28] (this effect is reversed in readers of Hebrew who read from right to left [29]).
102 Consistently, it has been shown that occluding the word just to the right of fixation, reduces
103 reading speed by 25–40 ms per word [11]. Interestingly, making the fixated word disappear
104 after 60 ms hardly impacts reading, whereas making the parafoveal word disappear after 60ms,
105 increases reading times substantially [30]. This finding provides strong evidence that fluent
106 reading relies on previewing. How deeply is upcoming parafoveal text previewed? There is
107 strong evidence for previewing at the sub-lexical level (e.g. **orthographic** and **phonological**
108 **processing**) [31, 32]. Using gaze-contingent boundary paradigms it was shown that fixation
109 times on the target word were reduced after it was primed in the parafovea by an
110 orthographically similar letter string (e.g. *sorp* priming *song*) compared to an unrelated
111 condition (*door* priming *song*) (e.g. [33, 34]). A similar effect has been found with respect to
112 phonological previewing using homophones [35, 36]. Previewing at the *lexical level* has been
113 investigated using sentences containing target words of low or high lexical frequency. Several
114 eye-tracking studies have found that pre-target fixation times are not modulated by the lexical
115 frequency of the target word, suggesting the absence of lexical previewing [37] (but see [38]).
116 We recently challenged this notion by combining eye-tracking with a **rapid frequency tagging**
117 paradigm. In this paradigm, we subliminally flickered the target words at 60 Hz during natural
118 reading. Using MEG it is then possible to measure the neuronal response at 60 Hz which is
119 specifically associated with the target word. This rapid frequency tagging approach allows for
120 assessing neuronal activity in the visual cortex on a fast time-scale [39]. We found that when
121 readers fixated on the pre-target words, there was a stronger tagging response, as measured by
122 MEG, when the target words were of low compared to high lexical frequency [40], which was
123 not reflected in the pre-target fixation durations. Additionally, faster readers showed stronger
124 parafoveal processing at the neuronal level, indicating that the amount of parafoveal processing
125 affects efficiency in processing. This finding provides novel neural support for previewing at
126 the lexical level, without necessarily impacting behavioral eye-movement measures. Another
127 controversial topic is whether there is previewing at the *semantic level*. This question has been

128 studied using boundary paradigms in which target words changed between semantically related
129 (e.g. *tune* to *song*) or unrelated words (*door* to *song*) when saccading to it. The target word
130 fixation times were not reduced following the semantic parafoveal priming indicating no
131 semantic preview benefits [41]. Surprisingly, when similar studies were conducted using
132 Chinese [42] and German [43], there was support for semantic previewing. In sum, while there
133 is evidence for previewing at the sub-lexical level, there are mixed findings regarding lexical
134 and semantic previewing.

135 136 **Pipelining as an alternative to serial or parallel processing**

137 Different mechanisms have been proposed to account for efficient visual processing, especially
138 during natural reading. It is strongly debated whether visual processing of foveal and
139 parafoveal words during reading is supported by a serial or a parallel mechanism [44].
140 Proponents of serial mechanisms argue that words are processed lexically one at a time (Fig
141 2a) [45, 46]. This does not preclude the processing of parafoveal words; however, this is
142 achieved by attention being allocated to parafoveal words after completing the lexical
143 processing of the fixated word (before moving the eyes). These principles are implemented in
144 the E-Z reader model, which can account for a large set of behavioural data [47, 48]. Serial
145 processing is challenged by researchers arguing for a parallel mechanism [49]. The core
146 argument is that words in the fovea and the parafovea are processed simultaneously, albeit in
147 a graded manner (Fig 2b), at the lexical, semantic and even syntactical level [38, 50, 51]. One
148 key argument is based on the phenomenon of word transpositions in reading. It has been shown
149 that readers have trouble classifying a sentence as syntactically incorrect if it can be corrected
150 through the transposition of two words ('You That Read Wrong Again!') [52].

151
152 The SWIFT and OB1 models for parallel processing have been developed to account for word
153 recognition in relation to saccade generation [53, 54]. However, the parallel mechanism faces
154 its own challenges. For instance, multiple objects need to be recognized simultaneously, which
155 may result in a bottleneck problem. The bottleneck problem has for instance been studied using
156 crowding paradigms in which the ability to recognize objects is impaired by clutter [55] and it
157 is likely to be a consequence of the hierarchical organization of the temporal lobe. In terms of
158 word recognition and lexical access, the bottleneck problem has been studied using behavioural
159 paradigms and fMRI [56, 57].

161 We here put forward a pipelining mechanism that incorporates elements from both models
162 while at the same time being compatible with the hierarchical organization of the visual system.
163 The pipelining process is implemented as a multiplexing mechanism, as explained in Fig 2c.
164 After orthographic processing of the fixated word *jumped* is completed, lexical processing is
165 initiated. This then allows for the parafoveal word *over* to be processed simultaneously at the
166 orthographic level. As such, serial processing occurs within a specific brain region, whereas
167 parallel processing occurs across different brain regions. The proposed pipelining scheme does
168 require a precise timing mechanism. As we explain below, neuronal oscillations are ideally
169 suited for coordinating the required temporal organization.

170

171 Some components of the E-Z Reader model might be compatible with a pipelining mechanism.
172 For instance, in both models there is a distinction between an early lower-level (orthographic)
173 stage and a later higher-level (semantic) stage of word processing, and both models allow for
174 substantial previewing of upcoming words. However, an important difference is that while
175 previewing or parafoveal processing in the E-Z Reader model is generally delayed until the
176 fixated word has been fully processed (unless the next word is very short, highly frequent,
177 and/or highly predictable), the pipelining mechanism allows for parafoveal processing to start
178 earlier in the cycle, alleviating the bottleneck problem.

179

180 **A mechanism supporting natural vision and reading by pipelining coordinated by alpha** 181 **oscillations**

182 We propose a pipelining mechanism that can be used to guide efficient visual exploration and
183 reading. We hypothesize that visual exploration relies on a process in which several objects are
184 processed simultaneously at different levels in the cortical hierarchy. Consider Figure 3A in
185 which the viewer fixates on the *woman*. The visual input propagates in the cortical hierarchy
186 in which features of increasing complexity are combined to semantically recognize the object
187 *woman* in inferior temporal (IT) cortex. While the participant fixates on the *woman*, the *dog* is
188 parafoveally previewed as a potential saccade goal. The previewing creates a bottleneck
189 problem due to the hierarchical nature of the visual system as two objects (e.g. *woman* and *dog*)
190 have to be recognized simultaneously. We propose that the bottleneck problem is solved by a
191 pipelining mechanism in which several objects are processed simultaneously but at different
192 levels in the cortical hierarchy (detailed in Figure 2 and Figure 3B). As a consequence, serial
193 processing occurs within a specific brain region, whereas parallel processing occurs across

194 different brain regions. This scheme also serves to coordinate the parafoveal previewing of the
195 *dog* which then will speed up the recognition time when the object eventually is fixated. The
196 pipelining mechanism requires tight temporal coordination. We propose that oscillations in the
197 8–13 Hz alpha band serve to organize visual presentations according to a **phase code** to support
198 parafoveal previewing and eventually guide the saccadic trajectory. In Figure 3 we have
199 assumed the **alpha oscillations** to be 12 Hz; however, is it possible to scale the mechanism to
200 operate with e.g. 10 Hz alpha oscillations; this will ease the temporal constraints slightly. The
201 saccades must be locked to the phase of the alpha oscillations in order for the processing to be
202 coordinated.

203 A similar mechanism might support natural reading (see Figure 4), with the exception that the
204 saccades are directed to the right. When the word *jumped* is fixated, this allows for *over* to be
205 previewed. As such, different words are processed simultaneously, but at different levels in the
206 hierarchy. As mentioned before, parafoveal previewing results in **lexico-semantic** recognition
207 to be reduced from about 150 ms to 110 ms. This provides extra time for evaluating the next
208 saccade goal and potentially skip a less informative word (e.g. *the*). The pipelining mechanism
209 is organized in a temporal code along the alpha cycle. Importantly, the saccades are locked to
210 the phase of the alpha oscillations in order to organize the processing and the visual input. In
211 short, we argue that efficient visual exploration and reading rely on parafoveal previewing, and
212 the bottleneck problem in the cortical hierarchy is resolved by a pipelining mechanism. The
213 multiplexed processing of fixated and previewed objects is coordinated by alpha oscillations.

214 A computational mechanism organized as a pipeline requires an intricate temporal organization
215 (Figure 3 and 2). The transfer of representations between levels across the hierarchy as well as
216 the sequential processing of multiple objects within the hierarchy must be coordinated. In the
217 example Figure 3, some of the visual features of the *boy* will propagate from V4 to face-
218 selective areas in IT. Likewise, the face-selective area will process the *boy* slightly earlier than
219 the *woman*. Based on recent findings, we propose that oscillatory coupling in the alpha band
220 serves to coordinate the information transfer between regions [58, 59].

221 While alpha oscillations for decades were thought to reflect idling or a state of rest [60], it is
222 now evident that they are involved in numerous cognitive processes [61-63]. One key insight
223 is that alpha oscillations are present during continuous visual processing (e.g. [64, 65]). Based
224 on human and animal data, the case has been made that theta and alpha oscillations reflect
225 pulses of inhibition that rhythmically interrupts neuronal firing [63, 66, 67]. These pulses might
226 be implemented by bouts of GABAergic interneuronal input. At the peak of an inhibitory pulse,

227 neurons are prevented from firing. As the inhibition ramps down over the cycle, the most
228 excitable neurons will fire first, then the somewhat less excited neurons and so forth. As such,
229 the pulses of inhibition implement a type of filter, ensuring that neuronal representations are
230 activated sequentially according to excitability [67, 68]. Representations of fixated objects are
231 associated with more neuronal excitability compared to parafoveal representations. This allows
232 foveal representations to overcome the pulsed inhibition earlier and thus activate earlier in the
233 alpha cycle. The sorting of different object representations according to excitability, creates the
234 phase-coded representations which are essential for the pipelining mechanism (Figure 3 and
235 2). While we have put forward an example with 2 objects in each cycle, the scheme could be
236 extended to 3-4 objects and to include more hierarchical levels. It is also interesting to consider
237 the framework in the light of mechanisms for working memory. We assume that a given word
238 or object is processed over multiple alpha cycles and the respective neuronal representations
239 must be maintained in this interval but not necessarily by sustained firing. Recent research has
240 suggested that working memory can be maintained by activity-silent states involving short-
241 term synaptic plasticity [69-71]. It would be of great interest to investigate if similar
242 mechanisms are at play when maintaining and integrating information over several alpha
243 cycles.

244 **Predictions and evidence in support of the pipelining mechanisms**

245 The mechanism outlined above rests on several assumptions, some of which are supported by
246 the literature while others need to be empirically tested. One assumption is that information in
247 the visual system is organized according to a phase-code coordinated by neuronal oscillations.
248 A second key assumption is that saccades are locked to the phase of the ongoing oscillations.
249 A third assumption is that the feed-forward flow is coordinated by phase-synchronization
250 between regions in the visual hierarchy in the ventral stream. We will here discuss the empirical
251 support for these assumptions.

252 **Prediction 1: phase-coding**

253 The proposed mechanism relies on organizing multiple representations in a phase code.
254 Recordings from hippocampal place cells in behaving rats have demonstrated the role of theta
255 oscillations (6 – 12 Hz) in organizing neuronal representations of space. The phenomenon of
256 *theta phase precession* shows that a given place cell fires late in the theta cycle as the rat enters
257 a place field. As the rat advances, the firing precesses to earlier theta phases. This finding is
258 best explained by a mechanism in which a sequence of spatial representations is ‘read out’
259 within a theta cycle [72, 73]. This phase-coding scheme is consistent with a pipelining

260 mechanism in which different representations along the path are sequentially processed at
261 different theta phases. There is an intriguing link between navigation and visual exploration:
262 both processes require that object representations are anchored to spatial representations in
263 order to represent the environment. It should also be mentioned that despite the different labels,
264 human alpha oscillations and rat theta oscillations overlap in frequency and might support
265 related functions [74].

266 Phase-coding with respect to neuronal oscillations has also been identified from intracranial
267 recordings in humans performing visual and working memory tasks. For instance, using human
268 intracranial data it was demonstrated that individual working memory representations are
269 represented at different phases of an 8 Hz rhythm [75]. Another intracranial study found that
270 different visual categories were encoded at different phases of the theta oscillations [76]. In
271 non-human primates, work based on intracranial recordings also reports on phase coding in the
272 visual system in various tasks [77-79].

273 While these reports are encouraging, the phase-coding scheme still needs to be investigated in
274 the context of visual exploration and reading. As outlined in Figures 2 and 3, we predict that
275 alpha-band oscillations coordinate the neuronal processing associated with saccadic visual
276 exploration. As a result, representationally-specific representations for fixated and upcoming
277 saccade goals should be coupled to the phase of the alpha oscillations. This could be tested by
278 MEG or EEG recordings in combination with eye-tracking in humans engaged in visual
279 exploration or reading tasks. It has been shown that different levels of representation associated
280 with upcoming information can be detected using multivariate approaches in MEG and EEG
281 [80, 81]. Therefore, the time-course of the representationally specific activation could be
282 identified by multivariate approaches and then be related to the phase of the ongoing alpha
283 oscillations [82]. As a complementary approach, rapid frequency tagging at different
284 frequencies (50 to 70 Hz) could be used to track several objects and the respective neuronal
285 signals would reflect the neuronal processing of the tagged objects [40]. This could also be
286 achieved using broad-band flicker and temporal response functions (see [83] for methodology).
287 Specifically, we predict that the fixated, as well as the parafoveal object (or word), would
288 become active at different phases of the alpha cycle (see Figure 3C and Figure 4B).

289

290 **Prediction 2: alpha oscillations linked to saccades**

291 A second key prediction is that saccades must be linked to the phase of alpha oscillations in
292 order to time the neuronal processing of the visual input. Studies in both humans and animals
293 have found intriguing links between the phase of alpha oscillations and saccades. A study based
294 on both MEG and intracranial human data showed that saccade onsets are locked to the phase
295 of ongoing alpha oscillations when viewing natural images [84]. Importantly, the degree of
296 locking predicted which pictures were later remembered as compared to forgotten. This
297 suggests that visual information impacts memory areas stronger when saccades are coordinated
298 by the phase of alpha oscillations. Using EEG, it was shown that the phase of pre-stimulus
299 alpha oscillations was associated with saccadic response latencies [85]. In another non-human
300 primate study, signals from the V4 receptive field of respectively current and future fixations
301 were coherent in the alpha band around the time of saccades [86]. Not only the allocation of
302 overt attention is related to the alpha phase, so is the allocation of covert attention. A recent
303 study in non-human primates reported on multi-electrode recording in the frontal-eye field that
304 allowed for decoding of the focus of the attentional spotlight. Importantly, the spotlight
305 explored the visual space at a 7-12 Hz alpha rhythm ('attentional saccades') [87]. Finally, there
306 is a more tentative link to explore between saccadic suppression and the conjectured pulses of
307 inhibition by alpha oscillations: recording sites in V4 associated with peripheral vision increase
308 in alpha power during saccades [88]. In short, strong evidence is accumulating in support of an
309 intimate connection between saccades and alpha oscillations. Further research is required to
310 establish a link between oscillations and saccades in relation to the proposed phase coding
311 scheme. Specifically, we predict that saccades would be locked to the alpha phase during visual
312 exploration and reading. Possibly a stronger locking would be associated with more efficient
313 visual processing.

314
315 While we have not addressed the putative role of microsaccades, this might eventually also be
316 relevant. It has been shown that microsaccades reflect the allocation of spatial attention [89]
317 [90], but surprisingly, microsaccades are mainly leftwards during reading [91]. EEG studies
318 have demonstrated that micro-saccades evoke activity in the alpha band [92] which warrant
319 further investigations.

320 321 **Prediction 3: Inter-regional phase-synchronization in the alpha band**

322 A third key requirement for the proposed scheme is that the feed-forward flow of the
323 representations is coordinated by phase-synchronization in the alpha band between regions in

324 the visual hierarchy. This would also support the exchange of phase-coded representations
325 between brain regions [58, 93]. In support, intracranial recordings in non-human primates have
326 shown that synchronization in the alpha band reflects forward communication between V4 and
327 interior temporal areas [94]. Human MEG recordings have demonstrated phase-
328 synchronization in the alpha band between regions in a larger visual hierarchy in participants
329 performing visual attention tasks [95]. We argue that the interregional phase-synchronization
330 in the alpha band could reflect the representational specific feed-forward flow [59, 93]. Other
331 support for temporal coordination by alpha oscillations stems from work demonstrating
332 travelling waves across regions in the alpha band [96], which even are task-dependent [97]. In
333 future work, it would be important to further uncover the phase-relationship between regions
334 in the visual hierarchy during visual exploration and reading tasks. This could be done using
335 MEG or intracranial recordings in patients undergoing presurgical evaluations.

337 **Concluding remarks**

338 We argue that the visual system must operate in a highly efficient manner to support visual
339 exploration and reading. The core issue is that the fixated object or word must be processed in
340 the same time interval as when the next saccade goal is planned. Given the bottleneck problem
341 in the visual hierarchy [46, 56], we propose that this is achieved by a pipelining mechanism
342 coordinating the processing of current and upcoming visual objects. Importantly, we propose
343 that neuronal oscillations in the alpha band serve to coordinate the pipelining mechanism that
344 is implemented by a phase-coding scheme in which different representations activate
345 sequentially along the phase of the alpha oscillations. Finally, to coordinate the visual input
346 with the neuronal processing, the mechanism also requires that saccades are locked to the phase
347 of the alpha oscillations. The proposed framework has resulted in a set of questions to be
348 addressed in future empirical investigations (see Outstanding Questions).

349 In sum, we have here presented a novel and testable framework for the neuronal mechanisms
350 supporting visual exploration and reading in relation to saccades. Crucially, neuronal
351 oscillations are required for organizing the visual representation as well as the timing of
352 saccades. Since the proposed mechanism provides a unified account for visual exploration and
353 reading, it also opens the door for future investigations aimed at understanding the neuronal
354 substrate associated with reading and visual disorders.

355 **Acknowledgements**

356 The work was supported by the following funding: James S. McDonnell Foundation,
1 357 Understanding Human Cognition Collaborative Award, grant number 220020448; the
2 358 Wellcome Trust Investigator Award in Science, grant number 207550; Biotechnology and
3 359 Biological Sciences Research Council, grant BB/R018723/1; Wolfson Research Merit Award
4 360 to Jensen, The Royal Society.
5
6

7 361 **Captions:**
8
9

10 362 **Figure 1. Temporal constraints during visual exploration and reading.** A) In this example
11 363 the participant first fixates on the *boy* and then saccades to the *woman* followed by the *dog*. B)
12 364 The timing of the visual exploration process. The visual object (the *boy*) arrives at 60 ms in V1
13 365 after which visual features are identified at about 85 ms [8]. Electrophysiological evidence
14 366 suggests that objects can be identified before 150 ms in object selective cortex [16-19]. While
15 367 this process is going on, the next saccadic decision must be made (*woman* or *dog*?) such that
16 368 the motor program can commence. It takes about 100 ms to initiate and execute the saccadic
17 369 motor command [1, 20]. Typically, a saccade is executed about 250 ms after the fixation onset.
18 370 As both the object identification as well as the saccade decision must be performed with 150
19 371 ms, this places serious computational demands on the visual system. C) Previewing by
20 372 parafoveal processing allows for speeding up visual processing. For instance, when fixating on
21 373 the *boy* the *woman* can be previewed. When the *woman* then is fixated the recognition can be
22 374 reduced to about 110 ms [1, 2, 11]. This has two important advantages: 1) it leaves about 40
23 375 ms for previewing the next saccade goal (the *dog*) and 2) the preview occurs sufficiently early
24 376 to impact the next saccade goal (e.g. to skip). D) A sentence is read by fixating on the words
25 377 sequentially. When fixating on the word *jumped* it must simultaneously be decided on whether
26 378 to saccade to *over*. E) The timing of the visual reading process. For instance, visual features of
27 379 the word *over* arrive at 60 ms in V1 after which the word-form is identified in the visual word-
28 380 form area (VWFA) at ~90 ms. There is electrophysiological evidence suggesting that lexical
29 381 recognition of the word is done within 150 ms by a network including the mid-fusiform gyrus
30 382 and possibly the anterior temporal lobe (ATL) [16-19]. While this process is going on, the next
31 383 saccade decision (*over*) must be made such that the saccadic motor program can be initiated.
32 384 Both the lexical identification as well as the saccade decision must be performed within 150
33 385 ms. F) Parafoveal previewing of a word (e.g. *over*) allows for reducing the lexico-semantic
34 386 identification upon fixation. As such a previewed fixated word could be recognized at 110 ms.
35 387 This has an important advantage: it leaves more time for previewing the next word and deciding
36 388 the next saccade goal before executing the saccade plan. For instance, a decision might be made
37 389 to skip highly common and/or predictable words.
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

390

- 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
- 28
- 29
- 30
- 31
- 32
- 33
- 34
- 35
- 36
- 37
- 38
- 39
- 40
- 41
- 42
- 43
- 44
- 45
- 46
- 47
- 48
- 49
- 50
- 51
- 52
- 53
- 54
- 55
- 56
- 57
- 58
- 59
- 60
- 61
- 62
- 63
- 64
- 65

391

392

Figure 2. Different models for reading. The eye symbols represent the fixations on the respective words. (a) Serial processing in which each word is processed one at a time. Note that the model still allows for parafoveal on foveal processing but only after the lexical processing of the fixated word is completed. (b) Parallel models for reading suggest that words in the fovea and parafovea are processed simultaneously even at the lexical level. This processing might be graded. (c) We propose that reading is supported by a pipelining mechanism that incorporates elements of both serial and parallel processing. In particular, foveal and parafoveal words (e.g. *jumped* and *over*) are processed simultaneously, however, at different cortical levels in the visual hierarchy.

Figure 3: A model for pipelining during visual exploration. A, B) The participant fixates on the *boy* and then saccades to the *woman* and finally to the *dog*. For simplicity, we assume two stages in which simple features (e.g. colour) are identified first (in V4), followed by object category recognition (in object selective cortex in the inferior temporal lobe). C) We hypothesize that the temporal organization supporting the pipelining mechanism is coordinated by oscillations in the alpha band. In this example, 12 Hz alpha oscillations can be considered pulses of inhibition repeated every 83 ms [67, 68]. Note, this is just an example and the mechanism would work with alpha oscillations at slower frequencies as well. At the peak of the alpha cycle, neuronal firing is inhibited. As the inhibition ramps down the most excitable representation will activate and so forth [68]. Consider time point $t = 0$ ms in which the participant moves the eyes from the *boy* and fixates on the *woman* (the black line on top indicates the horizontal gaze position). We assume that saccades are locked to the phase of the alpha oscillations such that the visual input of the *woman* arrives at the early down-going inhibitory slope of the alpha cycle at about $t = 85$ ms where simple visual features of the woman engage the visual occipital cortex (e.g. colours in V4). These feature representations are projected to object selective cortex for category identification by 110 ms. This fast category identification is made possible by the preview of the *woman* prior to the saccade which has primed the ‘semantic’ access. Importantly, the pipelining scheme allows the *woman* and the *dog* to be processed in the same cycle in a multiplex manner thus avoiding bottleneck problems. Specifically, the features of the *woman* and the *dog* are processed sequentially in V4 and

424 slightly later the respective object-categories are sequentially processed in object-selective
425 cortex. This scheme allows for a fast decision to be made to either saccade to the *dog* or hold
426 the saccade and preview another object as a potential target.

427
428
Figure 4: A model for pipelining during natural reading. A) For the sake of simplicity, two
430 stages of word recognition are assumed, namely orthographic identification in the visual word
431 form area (VWFA), followed by lexico-semantic access in an extended network including the
432 middle fusiform gyrus (MFG) and the anterior temporal lobe (ATL). B) We hypothesize that
433 the temporal organization supporting the pipelining mechanism is coordinated by oscillations
434 in the alpha band. The 12 Hz alpha oscillations can be considered pulses of inhibition repeated
435 every 83 ms [67, 68]. Consider time point $t = 0$ ms in which the subject saccades and fixates
436 on *jumped* (the black line on top indicates the horizontal gaze position). We assume that
437 saccades are locked to the phase of the alpha oscillations such that the visual input of *jumped*
438 arrives at the early down-going inhibitory slope of the alpha cycle at about 85 ms for
439 orthographic feature identification in the VWFA. The orthographic representations propagate
440 to the MFG or ATL for lexico-semantic identification by 110 ms. This fast lexico-semantic
441 process is made possible by the preview of *jumped* prior to the saccade that has primed the
442 lexical processing. Importantly, the pipeline scheme allows both *jumped* and *over* to be lexico-
443 semantically processed in the same cycle but at slightly different points in time, thus avoiding
444 a bottleneck problem in e.g. the MFG and ATL. During the fixation of *over* the word *the* is
445 previewed. Given that the word *the* carries little information and is highly frequent, a decision
446 to skip can be made. However, this is only possible if *over* has been previewed since this will
447 speed up the processing of *over* leaving more time to preview *the*. While this scheme for
448 simplicity only outlines orthographic and lexico-semantic processing, it could be further
449 developed by assigning the lexical and semantic processing to different alpha phases as well
450 as different regions (e.g. respectively the MFG and the ATL). It would also be of great interest
451 to extend the framework to include previewing of two (and not only) words.

453

1 454 **Glossary (450 words)**

2

3

4 455 **Alpha oscillations:** An 8-13 Hz neuronal rhythm associated with pulsed GABAergic inhibition
5
6 456 of neuronal firing.

7

8 457 **Lexical processing:** The process of retrieving a mental representation of a known word based
9
10 458 on orthographic or phonological representations.

11

12 459 **Multi-variate decoding:** An approach in which the multivariate signals from e.g. EEG/MEG
13
14 460 recordings are used to decode the distributed neuronal activity associated with a given feature
15
16 461 or object.

17

18 462 **Orthographic processing:** The process of identifying and combining graphemes to form
19
20 463 words.

21

22 464 **Phase-coding:** A temporal code in which different neuronal representations activate at
23
24 465 different phases of ongoing brain oscillations.

25

26 466 **Phonological processing:** The processing of sounds (e.g. phonemes) of language.

27

28
29 467 **Pipelining:** A serial mechanism in which multiple operations are performed in fast succession.

30

31 468 **Previewing:** The visual processing of parafoveal words or objects.

32

33
34 469 **Rapid frequency tagging:** A new approach in which parts of a visual scene are subliminally
35
36 470 flickered ('tagged') at high frequency (in the 50-70 Hz range). Recording the neuronal response
37
38 471 allows isolation of the neuronal activity associated with the tagged object.

39

40 472 **Semantics processing:** The process that binds together distributed information to form a single
41
42 473 concept associated with meaning.

43

44 474

45

46

47

48

49

50

51

52

53

54

55

56

57

58

59

60

61

62

63

64

65

475 **References (100 max)**

- 1 476 1. Reichle, E.D. and Reingold, E.M. (2013) Neurophysiological constraints on the eye-mind link. *Front*
- 2 477 *Hum Neurosci* 7, 361.
- 3 478 2. Becker, W. and Jurgens, R. (1979) An analysis of the saccadic system by means of double step
- 4 479 stimuli. *Vision Res* 19 (9), 967-83.
- 5 480 3. Yarbus, A.L. (1967) *Eye movements and vision*, Plenum Press.
- 6 481 4. Torralba, A. et al. (2006) Contextual guidance of eye movements and attention in real-world scenes:
- 7 482 the role of global features in object search. *Psychol Rev* 113 (4), 766-86.
- 8 483 5. White, A.L. et al. (2013) Adaptive deployment of spatial and feature-based attention before
- 9 484 saccades. *Vision Res* 85, 26-35.
- 10 485 6. Li, H.H. et al. (2019) Presaccadic attention improves or impairs performance by enhancing sensitivity
- 11 486 to higher spatial frequencies. *Sci Rep* 9 (1), 2659.
- 12 487 7. Liversedge, S.P. and Findlay, J.M. (2000) Saccadic eye movements and cognition. *Trends Cogn Sci* 4
- 13 488 (1), 6-14.
- 14 489 8. Schmolesky, M.T. et al. (1998) Signal timing across the macaque visual system. *J Neurophysiol* 79
- 15 490 (6), 3272-8.
- 16 491 9. Cichy, R.M. et al. (2014) Resolving human object recognition in space and time. *Nat Neurosci* 17 (3),
- 17 492 455-62.
- 18 493 10. Hung, C.P. et al. (2005) Fast readout of object identity from macaque inferior temporal cortex.
- 19 494 *Science* 310 (5749), 863-6.
- 20 495 11. Rayner, K. (1998) Eye movements in reading and information processing: 20 years of research.
- 21 496 *Psychol Bull* 124 (3), 372-422.
- 22 497 12. Carreiras, M. et al. (2014) The what, when, where, and how of visual word recognition. *Trends*
- 23 498 *Cogn Sci* 18 (2), 90-8.
- 24 499 13. Rayner, K. (2009) *Eye Movements in Reading: Models and Data*. *J Eye Mov Res* 2 (5), 1-10.
- 25 500 14. Reingold, E.M. et al. (2012) Direct lexical control of eye movements in reading: evidence from a
- 26 501 survival analysis of fixation durations. *Cogn Psychol* 65 (2), 177-206.
- 27 502 15. Sheridan, H. and Reichle, E.D. (2016) An Analysis of the Time Course of Lexical Processing During
- 28 503 Reading. *Cogn Sci* 40 (3), 522-53.
- 29 504 16. Sereno, S.C. et al. (2003) Context effects in word recognition: evidence for early interactive
- 30 505 processing. *Psychol Sci* 14 (4), 328-33.
- 31 506 17. Assadollahi, R. and Pulvermuller, F. (2003) Early influences of word length and frequency: a group
- 32 507 study using MEG. *Neuroreport* 14 (8), 1183-7.
- 33 508 18. Dambacher, M. et al. (2006) Frequency and predictability effects on event-related potentials
- 34 509 during reading. *Brain Res* 1084 (1), 89-103.
- 35 510 19. Woolnough, O. et al. (2021) Spatiotemporal dynamics of orthographic and lexical processing in the
- 36 511 ventral visual pathway. *Nat Hum Behav* 5 (3), 389-398.
- 37 512 20. Sereno, S.C. and Rayner, K. (2000) The when and where of reading in the brain. *Brain and Cognition*
- 38 513 42 (1), 78-81.
- 39 514 21. Pelli, D.G. and Tillman, K.A. (2008) The uncrowded window of object recognition. *Nat Neurosci* 11
- 40 515 (10), 1129-35.
- 41 516 22. Liversedge, S.P. et al. (2013) *The Oxford handbook of eye movements*, 2013 edition edn., Oxford
- 42 517 University Press.
- 43 518 23. Nuthmann, A. and Henderson, J.M. (2010) Object-based attentional selection in scene viewing. *J*
- 44 519 *Vis* 10 (8), 20.
- 45 520 24. Pajak, M. and Nuthmann, A. (2013) Object-based saccadic selection during scene perception:
- 46 521 evidence from viewing position effects. *J Vis* 13 (5).
- 47 522 25. Nuthmann, A. (2013) On the visual span during object search in real-world scenes. *Visual Cognition*
- 48 523 21 (7), 803-837.
- 49 524 26. LaPointe, M.R.P. and Milliken, B. (2016) Semantically incongruent objects attract eye gaze when
- 50 525 viewing scenes for change. *Visual Cognition* 24 (1), 63-77.

526 27. Coco, M.I. et al. (2020) Fixation-related Brain Potentials during Semantic Integration of Object-
 1 527 Scene Information. *Journal of Cognitive Neuroscience* 32 (4), 571-589.

2 528 28. Rayner, K. et al. (1980) Asymmetry of the effective visual field in reading. *Percept Psychophys* 27
 3 529 (6), 537-44.

4 530 29. Pollatsek, A. et al. (1981) Asymmetries in the perceptual span for Israeli readers. *Brain Lang* 14 (1),
 5 531 174-80.

6 532 30. Rayner, K. et al. (2006) Eye movements when reading disappearing text: the importance of the
 7 533 word to the right of fixation. *Vision Res* 46 (3), 310-23.

8 534 31. Starr, M.S. and Inhoff, A.W. (2004) Attention allocation to the right and left of a fixated word: Use
 9 535 of orthographic information from multiple words during reading. *European Journal of Cognitive*
 10 536 *Psychology* 16 (1-2), 203-225.

11 537 32. Pollatsek, A. et al. (1992) Phonological Codes Are Used in Integrating Information across Saccades
 12 538 in Word Identification and Reading. *Journal of Experimental Psychology-Human Perception and*
 13 539 *Performance* 18 (1), 148-162.

14 540 33. Rayner, K. et al. (1980) Integrating information across eye movements. *Cogn Psychol* 12 (2), 206-
 15 541 26.

16 542 34. McConkie, G.W. and Zola, D. (1979) Is visual information integrated across successive fixations in
 17 543 reading? *Percept Psychophys* 25 (3), 221-4.

18 544 35. Chace, K.H. et al. (2005) Eye movements and phonological parafoveal preview: effects of reading
 19 545 skill. *Can J Exp Psychol* 59 (3), 209-17.

20 546 36. Pollatsek, A. et al. (1992) Phonological codes are used in integrating information across saccades
 21 547 in word identification and reading. *J Exp Psychol Hum Percept Perform* 18 (1), 148-62.

22 548 37. Brothers, T. et al. (2017) Looking back on reading ahead: No evidence for lexical parafoveal-on-
 23 549 foveal effects. *Journal of Memory and Language* 96, 9-22.

24 550 38. Kliegl, R. et al. (2006) Tracking the mind during reading: the influence of past, present, and future
 25 551 words on fixation durations. *J Exp Psychol Gen* 135 (1), 12-35.

26 552 39. Zhigalov, A. et al. (2019) Probing cortical excitability using rapid frequency tagging. *Neuroimage*
 27 553 195, 59-66.

28 554 40. Pan, Y. et al. (accepted) Neural evidence for lexical parafoveal processing. *Nat Commun*.

29 555 41. Rayner, K. et al. (2003) On the processing of meaning from parafoveal vision during eye fixations
 30 556 in reading. *Mind's Eye: Cognitive and Applied Aspects of Eye Movement Research*, 213-234.

31 557 42. Yan, M. et al. (2009) Readers of Chinese extract semantic information from parafoveal words.
 32 558 *Psychon Bull Rev* 16 (3), 561-6.

33 559 43. Hohenstein, S. et al. (2010) Semantic preview benefit in eye movements during reading: A
 34 560 parafoveal fast-priming study. *J Exp Psychol Learn Mem Cogn* 36 (5), 1150-70.

35 561 44. Murray, W.S. et al. (2013) Serial and parallel processes in eye movement control: current
 36 562 controversies and future directions. *Q J Exp Psychol (Hove)* 66 (3), 417-28.

37 563 45. Reichle, E.D. et al. (2009) Encoding multiple words simultaneously in reading is implausible. *Trends*
 38 564 *Cogn Sci* 13 (3), 115-9.

39 565 46. White, A.L. et al. (2019) You Can't Recognize Two Words Simultaneously. *Trends Cogn Sci* 23 (10),
 40 566 812-814.

41 567 47. Reichle, E.D. et al. (2003) The E-Z reader model of eye-movement control in reading: comparisons
 42 568 to other models. *Behav Brain Sci* 26 (4), 445-76; discussion 477-526.

43 569 48. Reichle, E.D. et al. (2012) Eye movements in reading versus nonreading tasks: Using E-Z Reader to
 44 570 understand the role of word/stimulus familiarity. *Vis cogn* 20 (4-5), 360-390.

45 571 49. Snell, J. and Grainger, J. (2019) Readers Are Parallel Processors. *Trends Cogn Sci* 23 (7), 537-546.

46 572 50. Snell, J. et al. (2017) Evidence for simultaneous syntactic processing of multiple words during
 47 573 reading. *PLoS One* 12 (3), e0173720.

48 574 51. Kennedy, A. and Pynte, J. (2005) Parafoveal-on-foveal effects in normal reading. *Vision Res* 45 (2),
 49 575 153-68.

50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

576 52. Mirault, J. et al. (2018) You That Read Wrong Again! A Transposed-Word Effect in Grammaticality
1 577 Judgments. *Psychol Sci* 29 (12), 1922-1929.

2 578 53. Engbert, R. et al. (2005) SWIFT: a dynamical model of saccade generation during reading. *Psychol*
3 579 *Rev* 112 (4), 777-813.

4 580 54. Snell, J. et al. (2018) OB1-reader: A model of word recognition and eye movements in text reading.
5 581 *Psychol Rev* 125 (6), 969-984.

6 582 55. Levi, D.M. (2008) Crowding--an essential bottleneck for object recognition: a mini-review. *Vision*
7 583 *Res* 48 (5), 635-54.

8 584 56. White, A.L. et al. (2019) Parallel spatial channels converge at a bottleneck in anterior word-
9 585 selective cortex. *Proc Natl Acad Sci U S A* 116 (20), 10087-10096.

10 586 57. White, A.L. et al. (2020) Visual word recognition: Evidence for a serial bottleneck in lexical access.
11 587 *Atten Percept Psychophys* 82 (4), 2000-2017.

12 588 58. Jensen, O. (2001) Information transfer between rhythmically coupled networks: reading the
13 589 hippocampal phase code. *Neural Comput* 13 (12), 2743-61.

14 590 59. Bonnefond, M. et al. (2017) Communication between Brain Areas Based on Nested Oscillations.
15 591 *eNeuro* 4 (2).

16 592 60. Pfurtscheller, G. et al. (1996) Event-related synchronization (ERS) in the alpha band--an
17 593 electrophysiological correlate of cortical idling: a review. *Int J Psychophysiol* 24 (1-2), 39-46.

18 594 61. Van Diepen, R.M. et al. (2019) The functional role of alpha-band activity in attentional processing:
19 595 the current zeitgeist and future outlook. *Curr Opin Psychol* 29, 229-238.

20 596 62. Palva, S. and Palva, J.M. (2007) New vistas for alpha-frequency band oscillations. *Trends Neurosci*
21 597 *30* (4), 150-8.

22 598 63. Mazaheri, A. and Jensen, O. (2010) Rhythmic pulsing: linking ongoing brain activity with evoked
23 599 responses. *Front Hum Neurosci* 4, 177.

24 600 64. Wang, L. et al. (2018) Language Prediction Is Reflected by Coupling between Frontal Gamma and
25 601 Posterior Alpha Oscillations. *J Cogn Neurosci* 30 (3), 432-447.

26 602 65. VanRullen, R. and Macdonald, J.S. (2012) Perceptual echoes at 10 Hz in the human brain. *Curr Biol*
27 603 *22* (11), 995-9.

28 604 66. Mehta, M.R. et al. (2002) Role of experience and oscillations in transforming a rate code into a
29 605 temporal code. *Nature* 417 (6890), 741-6.

30 606 67. Jensen, O. et al. (2014) Temporal coding organized by coupled alpha and gamma oscillations
31 607 prioritize visual processing. *Trends Neurosci* 37 (7), 357-69.

32 608 68. Gips, B. et al. (2016) A biologically plausible mechanism for neuronal coding organized by the phase
33 609 of alpha oscillations. *Eur J Neurosci* 44 (4), 2147-61.

34 610 69. Wolff, M.J. et al. (2017) Dynamic hidden states underlying working-memory-guided behavior. *Nat*
35 611 *Neurosci* 20 (6), 864-871.

36 612 70. Li, J. et al. (2021) Temporally coherent perturbation of neural dynamics during retention alters
37 613 human multi-item working memory. *Prog Neurobiol* 201, 102023.

38 614 71. Masse, N.Y. et al. (2019) Circuit mechanisms for the maintenance and manipulation of information
39 615 in working memory. *Nat Neurosci* 22 (7), 1159-1167.

40 616 72. Skaggs, W.E. et al. (1996) Theta phase precession in hippocampal neuronal populations and the
41 617 compression of temporal sequences. *Hippocampus* 6 (2), 149-72.

42 618 73. Jensen, O. and Lisman, J.E. (1996) Hippocampal CA3 region predicts memory sequences:
43 619 accounting for the phase precession of place cells. *Learn Mem* 3 (2-3), 279-87.

44 620 74. Lisman, J.E. and Jensen, O. (2013) The theta-gamma neural code. *Neuron* 77 (6), 1002-16.

45 621 75. Bahramisharif, A. et al. (2018) Serial representation of items during working memory maintenance
46 622 at letter-selective cortical sites. *PLoS Biol* 16 (8), e2003805.

47 623 76. Watrous, A.J. et al. (2015) Phase-amplitude coupling supports phase coding in human ECoG. *Elife*
48 624 *4*.

49 625 77. Kayser, C. et al. (2012) Analysis of slow (theta) oscillations as a potential temporal reference frame
50 626 for information coding in sensory cortices. *PLoS Comput Biol* 8 (10), e1002717.

627 78. Montemurro, M.A. et al. (2008) Phase-of-firing coding of natural visual stimuli in primary visual
1 628 cortex. *Curr Biol* 18 (5), 375-80.

2 629 79. Turesson, H.K. et al. (2012) Category-selective phase coding in the superior temporal sulcus. *Proc*
3 630 *Natl Acad Sci U S A* 109 (47), 19438-43.

4 631 80. Wang, L. et al. (2018) Specific lexico-semantic predictions are associated with unique spatial and
5 632 temporal patterns of neural activity. *Elife* 7.

7 633 81. Wang, L. et al. (2020) Neural Evidence for the Prediction of Animacy Features during Language
8 634 Comprehension: Evidence from MEG and EEG Representational Similarity Analysis. *J Neurosci* 40 (16),
9 635 3278-3291.

10 636 82. van Es, M.W.J. et al. (2020) Phasic modulation of visual representations during sustained attention.
11 637 *Eur J Neurosci*.

13 638 83. Huang, Q. et al. (2018) Fast-backward replay of sequentially memorized items in humans. *Elife* 7.

14 639 84. Staudigl, T. et al. (2017) Saccades are phase-locked to alpha oscillations in the occipital and medial
15 640 temporal lobe during successful memory encoding. *PLoS Biol* 15 (12), e2003404.

16 641 85. Drewes, J. and VanRullen, R. (2011) This is the rhythm of your eyes: the phase of ongoing
17 642 electroencephalogram oscillations modulates saccadic reaction time. *J Neurosci* 31 (12), 4698-708.

18 643 86. Neupane, S. et al. (2017) Coherent alpha oscillations link current and future receptive fields during
19 644 saccades. *Proc Natl Acad Sci U S A* 114 (29), E5979-E5985.

20 645 87. Gaillard, C. et al. (2020) Prefrontal attentional saccades explore space rhythmically. *Nat Commun*
22 646 11 (1), 925.

23 647 88. Zanos, T.P. et al. (2016) Mechanisms of Saccadic Suppression in Primate Cortical Area V4. *J*
24 648 *Neurosci* 36 (35), 9227-39.

26 649 89. Hafed, Z.M. and Clark, J.J. (2002) Microsaccades as an overt measure of covert attention shifts.
27 650 *Vision Res* 42 (22), 2533-45.

28 651 90. Yuval-Greenberg, S. et al. (2014) Spontaneous microsaccades reflect shifts in covert attention. *J*
29 652 *Neurosci* 34 (41), 13693-700.

30 653 91. Bowers, N.R. and Poletti, M. (2017) Microsaccades during reading. *PLoS One* 12 (9), e0185180.

31 654 92. Dimigen, O. et al. (2009) Human microsaccade-related visual brain responses. *J Neurosci* 29 (39),
33 655 12321-31.

34 656 93. Quax, S. et al. (2017) Top-down control of cortical gamma-band communication via pulvinar
35 657 induced phase shifts in the alpha rhythm. *PLoS Comput Biol* 13 (5), e1005519.

36 658 94. Saalmann, Y.B. et al. (2012) The pulvinar regulates information transmission between cortical
37 659 areas based on attention demands. *Science* 337 (6095), 753-6.

39 660 95. Michalareas, G. et al. (2016) Alpha-Beta and Gamma Rhythms Subserve Feedback and
40 661 Feedforward Influences among Human Visual Cortical Areas. *Neuron* 89 (2), 384-97.

41 662 96. Bahramisharif, A. et al. (2013) Propagating neocortical gamma bursts are coordinated by traveling
42 663 alpha waves. *J Neurosci* 33 (48), 18849-54.

43 664 97. Pang, Z. et al. (2020) Turning the Stimulus On and Off Changes the Direction of alpha Traveling
44 665 Waves. *eNeuro* 7 (6).

46 666

Highlights

- Humans have a remarkable ability to efficiently acquire information from natural scenes and text by means of saccadic exploration.
- Visual exploration is surprisingly efficient given the temporal and spatial constraints imposed by the visual system. As such, both information from current fixations as well as upcoming parafoveal locations must be processed within a 150 ms time window.
- We propose a novel mechanism in which visual exploration and reading are supported by a pipelining in which serial processing occurs within a specific brain region, whereas parallel processing occurs across different brain regions.
- The timing of the pipelining mechanism is organized by alpha oscillations
- The pipelining mechanism predicts that fixated and parafoveal objects/words are represented at different phases of an alpha cycle.
- Consistent with the mechanism, data in humans and non-human primates point to a link between the timing of saccades and alpha oscillations.

Outstanding Questions

- How deeply are objects previewed before saccades are made to them? Are they previewed at the semantic level or maybe just in terms of features?
- How deeply are upcoming words previewed during reading before saccades are made to them? Are they primarily previewed at the sub-lexical (e.g. orthographical, phonological, orthographic) or also at the lexico-semantic level?
- What is the role of brain oscillations in visual exploration and reading? Recent studies have found cases where saccades are locked to the phase of alpha oscillations, but how general is this phenomenon in terms of task?
- Are different objects and words represented along the phase of oscillations in the alpha cycle during visual exploration and reading (akin to the coding scheme of place representation organized by theta oscillations observed in exploring rats)? This can be addressed using **multi-variate decoding** applied to MEG and EEG data in order to relate representational specific information to the phase of oscillations in the alpha band. Likewise, *rapid frequency tagging* can be used to investigate the allocation of visual resources already before a saccade in relation to ongoing brain oscillations.
- Is there a link between previewing abilities and reading disorders? For instance, impaired previewing during reading might account for some of the reading deficits observed in some types of developmental and acquired dyslexia. Can our proposed pipelining mechanism account for the impaired previewing?
- Recent evidence suggests that prediction plays an important role in natural vision and language comprehension. How do top-down predictions impact bottom-up parafoveal previewing during reading? Could it be that pre-activated representations and the representations of the previewed objects are encoded at the same phase of the alpha cycle?

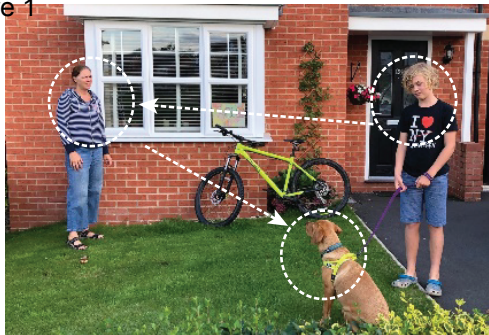
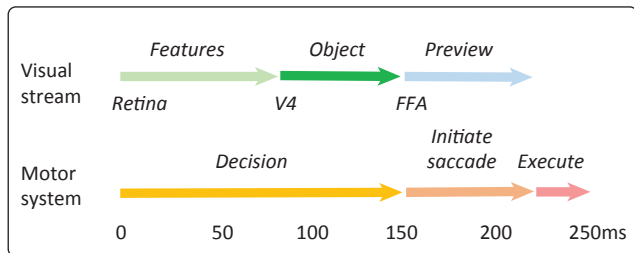
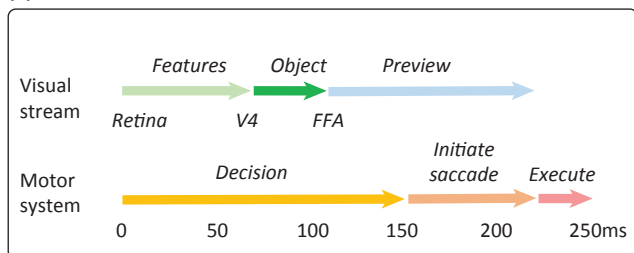
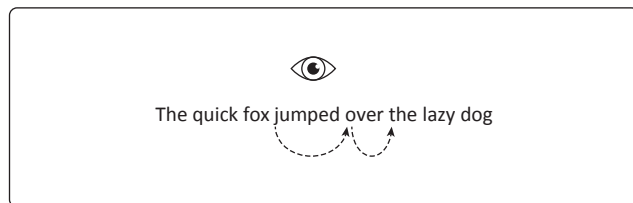
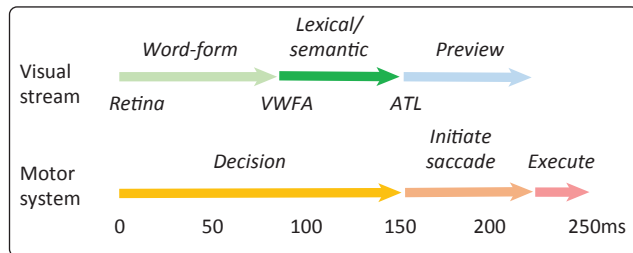
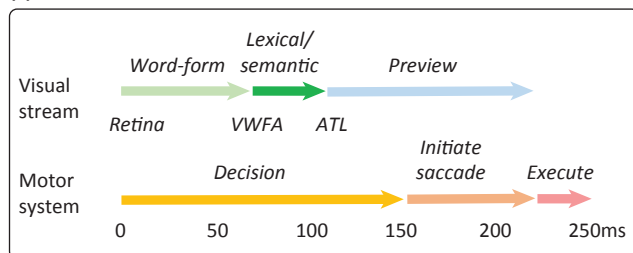
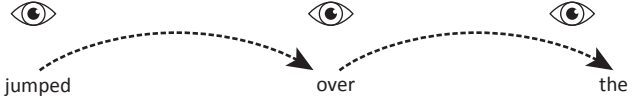
(Figure 1**(B) No preview benefit****(C) Preview benefit****(D)****(E) No preview benefit****(F) Preview benefit**

Figure 2

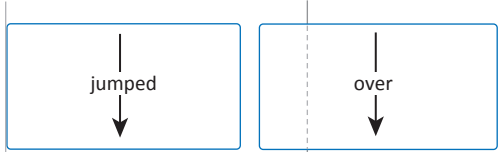


(A) Fixation

Orthography

Lexical

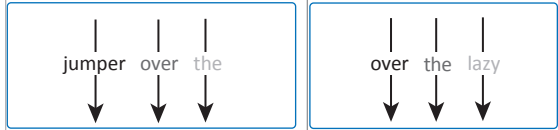
Semantic



(B) Orthography

Lexical

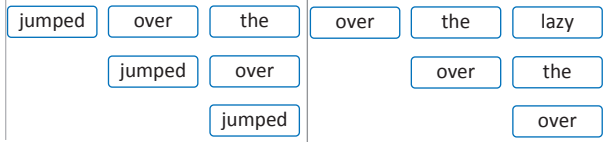
Semantic



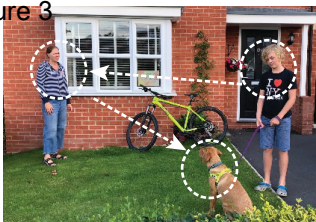
(C) Orthography

Lexical

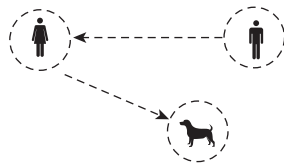
Semantic



(A) Figure 3



(B)



(C)

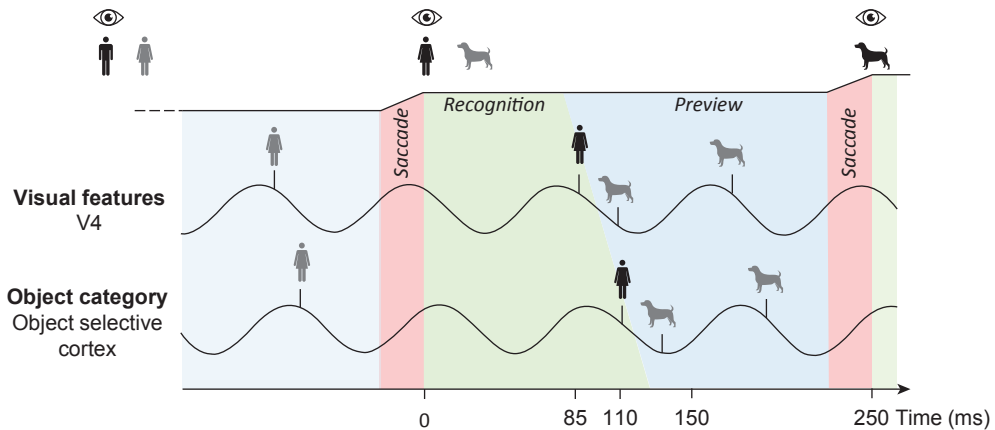


Figure 4

The quick brown fox jumped over the lazy dog The quick brown fox jumped over the lazy dog

