



HAL
open science

Visual attention modulates the transition from fine-grained, serial processing to coarser-grained, more parallel processing: A computational modeling study

Alexandra Steinhilber, Julien Diard, Emilie Ginestet, Sylviane Valdois

► To cite this version:

Alexandra Steinhilber, Julien Diard, Emilie Ginestet, Sylviane Valdois. Visual attention modulates the transition from fine-grained, serial processing to coarser-grained, more parallel processing: A computational modeling study. *Vision Research*, 2023, 207, pp.108211. 10.1016/j.visres.2023.108211 . hal-04084820

HAL Id: hal-04084820

<https://hal.science/hal-04084820>

Submitted on 19 Sep 2023

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.

Visual attention modulates the transition from fine-grained, serial processing to coarser-grained, more parallel processing: a computational modeling study

Alexandra Steinhilber¹, Julien Diard^{1,2}, Emilie Ginestet¹ and Sylviane Valdois¹

¹Laboratoire de Psychologie et NeuroCognition, Univ. Grenoble Alpes, Univ. Savoie Mont Blanc, CNRS, LPNC, 38000 Grenoble, France

²Corresponding author: julien.diard@univ-grenoble-alpes.fr

Abstract

During reading acquisition, beginning readers transition from serial to more parallel processing. The acquisition of word specific knowledge through orthographic learning is critical for this transition. However, the processes by which orthographic representations are acquired and fine-tuned as learning progresses are not well understood. Our aim was to explore the role of visual attention in this transition through computational modeling. We used the BRAID-Learn model, a Bayesian model of visual word recognition, to simulate the orthographic learning of 700 4-to 10-letter English known words and novel words, presented 5 times each to the model. The visual attention quantity available for letter identification was manipulated in the simulations to assess its influence on the learning process. We measured the overall processing time and number of attentional fixations simulated by the model across exposures and their impact on two markers of serial processing, the lexicality and length effects, depending on visual attention quantity. Results showed that the two lexicality and length effects were modulated by visual attention quantity. The quantity of visual attention available for processing further modulated novel word orthographic learning and the evolution of the length effect on processing time and number of attentional fixations across repeated exposures to novel words. The simulated patterns are consistent with behavioral data and the developmental trajectories reported during reading acquisition. Overall, the model predicts that the efficacy of orthographic learning depends on visual attention quantity and that visual attention may be critical to explain the transition from serial to more parallel processing.

Keywords: Visual attention, Bayesian modeling, Length effect, Orthographic learning

1. Introduction

1.1. Theoretical background

During learning to read, children move from slow serial processing to faster, more parallel word recognition (Castles et al., 2018). This developmental trajectory was initially conceptualized as reflecting successive stages in reading acquisition (Frith, 1985). However, the self-teaching theory (Share, 1995, 1999; Share and Shalev, 2004) proposed to replace this stage-based model by an item-based model according to which the transition from serial letter-by-letter to more parallel processing would apply at the level of each individual item word. According to this theory, the first time the child encounters a new printed word, this word would be serially processed through phonological recoding (i.e., translation of each orthographic unit into its spoken form). When phonological recoding is successful, then the input orthographic information can

be memorized, leading to enriching the reader's word-specific orthographic knowledge. Although some orthographic learning was demonstrated following a single encounter with the novel word, additional encounters contribute to shape well-specified word-specific orthographic representations (Bowey and Muller, 2005; Nation et al., 2007; Pellicer-Sanchez, 2016; Share and Shalev, 2004). The acquisition of new orthographic representations during reading (referred to as "orthographic learning" hereafter) allows fast recognition of previously encountered words, which is the hallmark of expert reading.

The self-teaching hypothesis is not age-specific. Most of the printed words beginning readers are exposed to are new words for them, which increases the probability of orthographic learning as soon as they have enough knowledge about print-to-sound mapping. However, readers are likely to be exposed to new words throughout their lifespan, so that orthographic learning

38 through self-teaching is observed in both beginning and
39 skilled readers (Bowers et al., 2005; Joseph and Nation,
40 2018; Joseph et al., 2014; Manis, 1985; Pagan and Na-
41 tion, 2019). Interestingly, the capacity to build-up new
42 words' orthographic knowledge across repeated expo-
43 sures may be as efficient in beginning as in more ad-
44 vanced readers (van Viersen et al., 2022), suggesting
45 that the same mechanisms are involved regardless of
46 reading practice.

47 Orthographic learning is characterized by a reduc-
48 tion of both the word length effect (i.e., additional pro-
49 cessing cost for longer words) and the lexicality effect
50 (i.e., differences in processing between unknown vs.
51 known words) in reading. Length effect on word reading
52 speed decreases with reading expertise and the develop-
53 ment of orthographic knowledge (Marinelli et al., 2016;
54 Provazza et al., 2019; Zoccolotti et al., 2005). This is
55 accompanied by changes in eye movements. Gaze dura-
56 tion and the probability of refixations is less influenced
57 by word length in more advanced readers (Joseph et al.,
58 2009; Rayner, 1998). A decrease of the length effect
59 on reading times was also reported with repeated ex-
60 posure to novel words in tasks of orthographic learn-
61 ing (Suarez-coalla et al., 2016). At fixed length, online
62 measures of eye movements across repeated exposures
63 to novel words and known words revealed larger learn-
64 ing effects for novel words (Ginestet et al., 2020; van
65 Viersen et al., 2022). A larger decrease in gaze duration
66 and fixation number across exposures was reported for
67 novel words, showing a reduction of the lexicality effect
68 with learning.

69 The importance of orthographic learning in the tran-
70 sition from novice to expert reading is now well es-
71 tablished. However, much is still unknown about the
72 mechanisms involved in orthographic learning. Ac-
73 cording to the self-teaching theory, phonological re-
74 coding is the primary mechanism by which an ortho-
75 graphic representation is acquired (Share, 1995, 1999;
76 Share and Shalev, 2004). The models of reading ac-
77 quisition that implement the self-teaching mechanisms
78 (Perry et al., 2019; Pritchard et al., 2018; Ziegler et al.,
79 2014), assume that phonological recoding relies on the
80 mapping of graphemes onto phonemes. Knowledge
81 about grapheme-to-phoneme mapping allows generat-
82 ing phonemic sequences that can trigger the activation
83 of known spoken words in phonological memory. When
84 an existing phonological word is sufficiently activated,
85 then an orthographic representation is set up in long-
86 term memory which is connected to the word phonolog-
87 ical representation and its meaning. Simulations within
88 these computational models have shown that most novel
89 words could be successfully learned through phonolog-

90 ical recoding (Perry et al., 2019; Pritchard et al., 2018;
91 Ziegler et al., 2014). In contrast, the role of visual pro-
92 cessing in orthographic learning is minimized in the
93 self-teaching theory (Share, 1999) and computational
94 modeling suggests that orthographic learning is more
95 sensitive to phonological than visual deficits (Perry
96 et al., 2019; Ziegler et al., 2014). However, these mod-
97 els make a number of simplifying assumptions about
98 the mechanisms of visuo-orthographic processing and
99 orthographic memorization. First, they do not imple-
100 ment the mechanisms of visual acuity, lateral interfer-
101 ence and visual attention that are known to modulate let-
102 ter identity processing within strings (Pelli et al., 2007;
103 Waechter et al., 2011) but rather postulate that accu-
104 rate identity information is immediately available for all
105 the letters within the input string. Second, they assume
106 that the word complete orthographic representation is
107 acquired in a "one-shot" manner, after a single expo-
108 sure (Perry et al., 2019; Pritchard et al., 2018; Ziegler
109 et al., 2014). This would predict an abrupt shift from
110 serial to parallel processing at the item level after a sin-
111 gle exposure, which contrasts with behavioral evidence
112 that successive exposures to words gradually shape or-
113 thographic representations (Ginestet et al., 2020; Joseph
114 et al., 2014; Nation et al., 2007; Pagan and Nation, 2019;
115 Pellicer-Sanchez, 2016; Suárez-Coalla et al., 2014).

116 Despite the importance of phonological recoding
117 in reading acquisition, there is behavioral evidence
118 that phonological processing cannot be the sole mech-
119 anism involved in the development of orthographic
120 knowledge. Self-teaching studies on typical read-
121 ers have shown that successful phonological recod-
122 ing only weakly predicted orthographic learning at the
123 item level, suggesting that other mechanisms were fur-
124 ther involved (Bosse et al., 2015; Cunningham, 2006;
125 Cunningham et al., 2002; Nation et al., 2007; Tucker
126 et al., 2016). The dissociations reported in developmen-
127 tal dyslexia between word-specific orthographic knowl-
128 edge and phonological recoding lead to the same con-
129 clusion, showing that good orthographic knowledge
130 might develop despite very poor phonological recod-
131 ing skills while, conversely, good phonological recod-
132 ing skills provided no guarantee of good orthographic
133 knowledge acquisition (Castles, 1996, 2006; Howard,
134 1996; Valdois et al., 2011, 2003). Furthermore, demon-
135 strations that humans can acquire orthographic knowl-
136 edge from artificial scripts that do not have any connec-
137 tion to phonology (Chetail, 2017; Lelonekiewicz et al.,
138 2020), and that nonhuman animals can acquire knowl-
139 edge about printed words without any language or
140 phonological skills (Grainger et al., 2012; Scarf et al.,
141 2016), suggest less phonological dependency in the de-

142 velopment of orthographic knowledge than currently 194
143 postulated. 195

144 Indeed, it has been shown that orthographic learn- 196
145 ing was facilitated when more visual information on the 197
146 input letter-string was simultaneously available during 198
147 the learning phase (Bosse et al., 2015). This suggests 199
148 that the mechanisms involved in visuo-orthographic 200
149 processing may represent additional components that 201
150 contribute to orthographic learning, independently of 202
151 phonological skills. Some insights on these mecha- 203
152 nisms comes from studies on the length effect in read- 204
153 ing (Barton et al., 2014). The fact that longer words (fam- 205
154 ilar or not) are fixated for longer than shorter words 206
155 and have a higher probability to be refixated (Hautala 207
156 et al., 2011; Kliegl et al., 2004; Loberg et al., 2019; 208
157 Lowell et al., 2014; McDonald, 2006; Rayner et al., 209
158 1996; Vitu et al., 1990) was interpreted as following 210
159 from the fact that more letters would fall in regions 211
160 of poorer visual acuity in longer words, thereby reduc- 212
161 ing the probability of successful identification (Engbert 213
162 et al., 2002; Reichle et al., 2003). However, length ef- 214
163 fects on eye movements have also been reported when 215
164 words were equated for their spatial extent, so that vi- 216
165 sual acuity decline was similar for all words whatever 217
166 their length (Hautala et al., 2011; McDonald, 2006). Ev- 218
167 idence for a length effect beyond the influence of vi- 219
168 sual acuity was interpreted as potentially reflecting dif- 220
169 ferential crowding effects, assuming that more letters 221
170 suffered from crowding (i.e., interference from adja- 222
171 cent letters) in longer than in shorter words (Hautala 223
172 et al., 2011; McDonald, 2006). However, visual acu- 224
173 ity and crowding can hardly account for the evolution 225
174 of eye movement behavior in condition of orthographic 226
175 learning, in which readers are repeatedly exposed to the 227
176 same set of words (at fixed length) (Ginestet et al., 2020; 228
177 Joseph and Nation, 2018; Joseph et al., 2014; Pagan and 229
178 Nation, 2019; Pellicer-Sanchez, 2016).

179 Visual attention is a third mechanism involved in 230
180 letter-string processing that might further affect ortho- 231
181 graphic learning. Behavioral studies have mainly fo- 232
182 cused on the visual attention span (VAS), a measure 233
183 of multielement parallel processing (Frey and Bosse, 234
184 2018; Valdois, 2022). VAS is known to relate to 235
185 reading acquisition (Valdois et al., 2019) and children 236
186 with higher VAS show higher reading fluency (Bosse 237
187 and Valdois, 2009) and higher orthographic knowledge 238
188 (Niolaki et al., 2020). By reference to the “Theory of 239
189 Visual attention” (Bundesen, 1990), VAS was found to 240
190 reflect the amount of visual attention available for mul- 241
191 tielement processing (Bogon et al., 2014; Dubois et al., 242
192 2010; Lobier et al., 2013). Neuroimaging studies re- 243
193 vealed that VAS related to the activation of the superior

parietal lobules within the dorsal attentional network 194
(Lobier et al., 2012; Peyrin et al., 2011; Reilhac et al., 195
2013). Only a few behavioral studies have examined 196
whether VAS was involved in orthographic learning. In 197
an experiment conducted in adults, Ginestet et al. (2020) 198
showed that orthographic learning and eye movement 199
patterns across exposures were modulated by VAS. Us- 200
ing a self-teaching paradigm without eye-movement 201
monitoring in children, Marinelli et al. (2020) showed 202
that VAS contributed to promote orthographic learning. 203

204 Interestingly, VAS was further described as relating 205
to the length effect in reading. Lower length effects 206
on word and pseudo-word reading latencies were re- 207
ported in individuals with higher VAS (van den Boer 208
et al., 2013) and exaggerated length effects were found 209
in individuals suffering from a VAS reduction (Juphard 210
et al., 2004; Valdois et al., 2011, 2003). In addition to 211
visual acuity and crowding, these findings suggest that 212
visual attention might be involved not only in the way 213
words are processed (i.e., in a strict serial or more par- 214
allel manner), but further in the capacity to acquire new 215
orthographic representations.

1.2. *The present study*

216 The main contribution of the present study was to 217
investigate the role of visual attention in orthographic 218
learning using a modeling approach. For this purpose, 219
we started from BRAID, a word recognition model that 220
implements the three mechanisms of visual attention, 221
visual acuity and lateral interference that are known to 222
affect letter identification within strings (Ginestet et al., 223
2019; Phenix, 2018; Phénix et al., 2018; Saghiran et al., 224
2020). In BRAID, the spatial distribution of visual at- 225
tention was modeled by a Gaussian probability distri- 226
bution, so that the letters near the focus (i.e., peak) of 227
attention were better recognized while the number of 228
letters that were allocated attention was dependent on 229
attention dispersion. Computational studies have shown 230
that variations in visual attention dispersion modulated 231
word recognition (Valdois et al., 2021a) and the word 232
length effect in tasks of lexical decision, naming and 233
progressive demasking (Ginestet et al., 2019; Saghiran 234
et al., 2020). The initial word recognition model was 235
extended in BRAID-Learn, a model of orthographic 236
learning (Ginestet, 2019; Ginestet et al., 2022). The 237
model incorporates a mechanism of visual attention ex- 238
ploration that optimizes the gain of information on letter 239
identity within the input string over time through modu- 240
lation of the two parameters of attentional focus location 241
and attention dispersion.

242 Ginestet et al. (2022) showed that BRAID-Learn suc- 243
244 cessfully simulated the evolution of eye-movement pat-

245 terns across repeated exposure to novel words by skilled 297
246 readers. This was mainly due to the interaction of 298
247 bottom-up sensory information modulated by visual 299
248 attention and top-down lexical feedback from the newly 300
249 acquired orthographic representation. However, the 301
250 study focused on words of fixed length and attention 302
251 quantity in the model was defined by its default value, 303
252 thus remaining constant through simulations. 304

253 Our purpose in the present study was to provide a 305
254 more plausible implementation of visual attention pro- 306
255 cessing in BRAID-Learn. Indeed, behavioral stud- 307
256 ies have shown that VAS increased with age dur- 308
257 ing childhood (from first to fifth grade) (Bosse and 309
258 Valdois, 2009) and that inter-individual variations in 310
259 VAS accounted for differences in orthographic learn- 311
260 ing (Ginestet et al., 2020). As VAS reflects the amount 312
261 of visual attention available for processing (Valdois, 313
262 2022), this suggests that a plausible model of ortho- 314
263 graphic learning should be able to simulate the conse- 315
264 quences of variations in visual attention quantity on pro- 316
265 cessing. Our main contribution was thus to introduce 317
266 a new visual attention quantity parameter in the model
267 and examine the effect of attention quantity variations
268 on orthographic learning through simulations. 318

269 Second, despite behavioral evidence that the length
270 effect on word and pseudo-word reading decreases with
271 reading expertise (Marinelli et al., 2016; Provasza et al.,
272 2019; Zoccolotti et al., 2005), evidence is lacking on
273 the evolution of length effects over repeated exposure
274 to known or novel words in condition of orthographic
275 learning. To fill this gap and provide new insights for
276 future behavioral studies, we examined the model’s pre-
277 dictions depending on the attention quantity available
278 for processing when repeatedly exposed to known or
279 novel words that varied in length. We used the model
280 as an experimental substitute to study the length effect
281 all other factors otherwise equal. For this purpose, a
282 single set of words was considered as known words in
283 a first series of simulations, in which the target words’
284 orthographic information was part of the model’s word
285 knowledge, but as novel words in a second series of sim-
286 ulations conducted after removing target words’ ortho-
287 graphic knowledge from the model database. 319

288 Assuming that higher visual attention quantity would
289 allow the model to accurately identify more letters si-
290 multaneously, we expected longer stimuli to be more
291 proficiently processed as attention quantity increases.
292 More proficient processing was expected to result in
293 shorter processing time (i.e., fewer iterations) and a
294 smaller number of attentional fixations during the visuo-
295 attentional exploration of the input word. Novel words
296 that do not benefit from top-down lexical knowledge at

the first exposure, would be processed less efficiently
than known words; moreover this difference would be
magnified with low attention quantity. However, or-
thographic learning being initiated at the first exposure,
novel word processing would improve across exposures
due to increasingly strengthened top-down support from
the newly acquired orthographic representation of the
target novel word. Assuming that higher attention quan-
tity allows processing more letters efficiently, ortho-
graphic knowledge acquisition would be more effective
at each exposure, leading to more proficient learning of
the novel word orthographic representation. This would
also result in a stronger length effect decrease, both on
processing times and number of attentional fixations,
across exposures as visual attention quantity is higher.

The rest of this paper is structured as follows. First,
we describe the BRAID-Learn model, with a particu-
lar focus on the visual-attention component. Second,
we detail the material and procedure used in the exper-
iment. Third, we present the simulation results, which
we discuss and relate to behavioral data.

2. The BRAID-Learn model

2.1. General outline of the model

The BRAID-Learn model shares the core of its ar-
chitecture with the three-layer architecture used, among
others, by the classical Interactive Activation model
(IA; McClelland and Rumelhart, 1981). It also fea-
tures an additional, original layer modeling visual at-
tention, along with mechanisms for orthographic learn-
ing. The resulting architecture is shown in Figure 1.
The BRAID-Learn model is a hierarchical, probabilistic
model, defined by a joint probability distribution over
its variables. As it is not relevant for the scope of the
current study, and as completely defining the model re-
quires space, we do not describe entirely its mathemat-
ical definition or its resulting properties here. However,
they can be found elsewhere (Ginestet, 2019; Phenix,
2018). Instead, in this section, we provide the necessary
elements to detail how orthographic learning processes
are implemented, and how visuo-attentional properties
affect the learning process.

The model includes four submodels. The letter sen-
sory submodel focuses on low-level mechanisms in-
volved in letter identification within the input string.
Letter identification at this level is modulated by inter-
letter visual similarity, implemented through a let-
ter confusion matrix adapted from experimental data
(Townsend, 1971) and by two mechanisms of visual

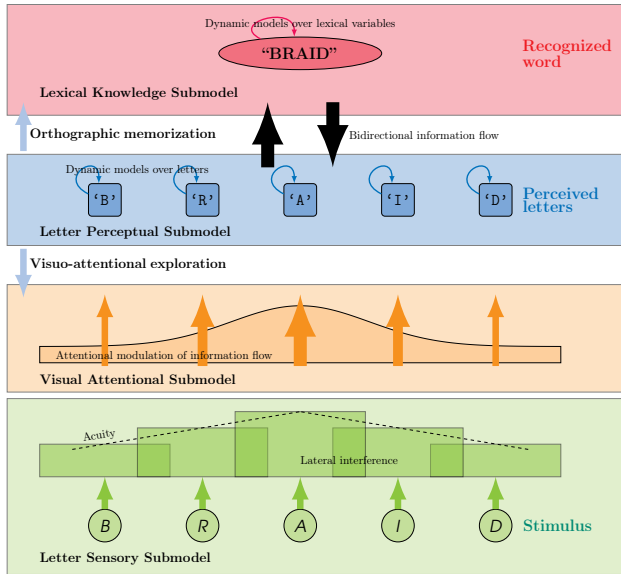


Fig. 1: Conceptual representation of the BRAID-Learn model. The four submodels are represented as colored blocks, and arrows represent information flow or specific processes (blue arrows). This structure is illustrated here on a 5-letter stimulus. See text for details.

acuity and lateral interference. The acuity gradient penalizes letter identification proportionally to the distance of the letter to gaze position. Letter identification is further affected by interference from neighboring letters so that inner letters suffer more interference than outer letters.

In the letter perceptual submodel, evidence about the identity of letters accumulates over time, to build dynamically evolving perceptual representations of the letters in the input string. These perceptual representations receive sensory information from the letter sensory submodel, in a bottom-up manner. They are further influenced by top-down information from the lexical knowledge submodel, so that identity information accumulates faster at the perceptual level for letters that belong to previously known words. (Note that, in the context of the current orthographic learning experiments, top-down knowledge about gradually improving orthographic traces is facilitatory. However, this is not a general property of the model. When top-down information from lexical knowledge is incongruent with the stimulus letters, for instance in priming simulations, it can slow down letter perception.)

The lexical knowledge submodel is configured to represent the spellings of a large database of words. The current simulations were run using a dataset of 79,673 English words, taken from the English Lexicon Project (Balota et al., 2007). The submodel further includes a

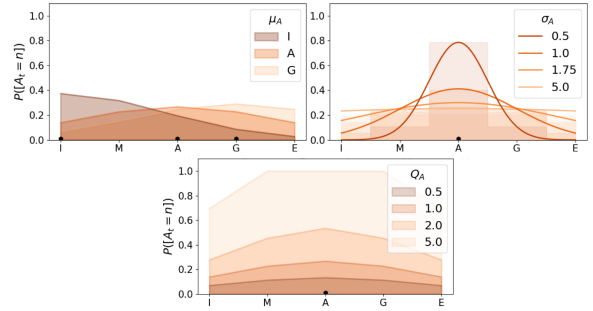


Fig. 2: Illustration of visuo-attentional distributions on the 5-letter input word “IMAGE”. Top left: attention distribution for a few values of parameter μ_A^t , which defines the position of attentional focus at time t . Top right: attention distribution for a few values of parameter σ_A^t , which defines attentional dispersion. Bottom: attention distribution for a few values of parameter Q_A , which defines the total attention quantity available for processing.

mechanism that evaluates lexical membership and allows determining whether the input stimulus is a known word or a novel word.

BRAID-Learn further includes a visuo-attentional submodel that controls the flow of information from the sensory to the perceptual submodel. Given the key role of visual attention in letter identity processing and orthographic learning, this submodel is described in more detail below.

2.2. The visuo-attentional submodel

The visuo-attentional submodel acts as a filter between the sensory and the perceptual submodels. Its main element is a Normal probability distribution, noted $P(A^t | \mu_A^t, \sigma_A^t)$, whose parameters μ_A^t and σ_A^t describe how visual attention is spatially distributed over the input letter string: the μ_A^t parameter represents the position of the focus of visual attention at time t (see Figure 2, top left), whereas the standard deviation σ_A^t parameter characterizes visual attention dispersion (see Figure 2, top right). Each letter in the input letter string, and therefore each position of the stimulus, is allocated a certain amount of visual attention, defined by this probability distribution. The amount of visual attention allocated to each position defines the amount of sensory evidence propagated from the sensory to the perceptual submodel. Due to the shape of the Gaussian distribution, less evidence on letter identity accumulates in the perceptual submodel when the distance from the attentional focus increases (Figure 2, top left). Letter identity processing is further modulated by visual attention dispersion. The smaller the attentional dispersion, the more attention is concentrated around the attentional fo-

405 cus, favoring efficient processing of a few letters, to the
 406 detriment of the others. The quality of perceptual repre-
 407 sentations is thus strongly modulated by the parameters
 408 of visual attention.

409 In previous simulations using either BRAID or an-
 410 terior versions of the BRAID-Learn model, the total
 411 amount of attentional resources available for process-
 412 ing was implicitly equal to 1 (its default value). In the
 413 context of the present study, we have defined the pa-
 414 rameter Q_A to explicitly represent the attention quan-
 415 tity. It is a multiplicative coefficient applied to the distri-
 416 bution of attention with the precaution that the amount
 417 of attention allocated to each position cannot exceed 1.
 418 Whatever the values of μ'_A and σ'_A , the higher Q_A ,
 419 the more attention is available for the processing of the at-
 420 tended letters, so that perceptual representations accu-
 421 mulate more identity evidence on these letters at each
 422 time-step, resulting, overall, in faster processing. The
 423 effect of parameter Q_A on the attention value at each
 424 position is illustrated on Figure 2 (bottom plot).

425 2.3. Orthographic learning in BRAID-Learn

426 In the model, orthographic learning consists in the
 427 transfer of letter identity information from the percep-
 428 tual submodel to the lexical knowledge submodel. As a
 429 result, orthographic learning is more efficient when per-
 430 ceptual information is of higher quality (i.e., providing
 431 enough information on letter identity at each position).

432 For the purpose of the current simulations, we con-
 433 sider that the model is given a task, in which the stim-
 434 ulus must be freely explored at each exposure, with
 435 no time-limit, until getting a precise enough percep-
 436 tual representation of the input letter string. At the end
 437 of each exposure, the lexical membership mechanism
 438 evaluates whether perceptual information corresponds
 439 to a known word, by comparing the perceived letters to
 440 known words' letters. If this is the case, then the exist-
 441 ing orthographic trace of the most likely word (a word
 442 recognition process, not detailed here, also proceeds in
 443 parallel) is updated by combining it with the perceptual
 444 representation of letters. Let us write (using a simplified
 445 notation) $P(P | S)$ the probability distribution about let-
 446 ters P given the stimulus letter sequence S , computed at
 447 the end of the exposure (i.e., the perceptual representa-
 448 tion of letters), $P(L_n | [W = w])$ the probability distri-
 449 bution over letters for word w in the set of known words
 450 W , after n exposures (i.e., the orthographic trace of word
 451 w), θ_n the learning rate after n exposures (it decreases
 452 exponentially across exposures), and finally U the uni-
 453 form distribution over the letter space. The probability
 454 distribution of the updated orthographic representation

455 after $n + 1$ exposures is as follows:

$$456 P(L_{n+1} | [W = w]) \propto$$

$$457 P(L_n | [W = w]) \times \left(\theta_n \times P(P | S) + (1 - \theta_n) \times U \right).$$

458 If, on the contrary, the perceptual information does
 459 not correspond to any word in the lexicon, then, a new
 460 orthographic trace is created. This trace is initialized
 461 with the perceptual representation of letters at the end
 462 of the first exposure. At each subsequent encounter with
 463 the “novel” word, the corresponding orthographic trace
 464 is gradually reinforced. Orthographic learning is said to
 465 be successful when the trace of an already encountered
 466 word is updated at subsequent encounters or when a new
 467 trace is created for a novel word at the first encounter.

468 The influence of lexical feedback on letter perception
 469 in the model is driven by lexical membership evalua-
 470 tion, so that the more likely the stimulus is to be a word,
 471 the stronger the lexical feedback. As a result, grad-
 472 ual strengthening of the orthographic trace makes novel
 473 word processing more and more efficient across expo-
 474 sures. A more detailed description of the mechanisms
 475 of lexical feedback and trace creation and updating can
 be found elsewhere (Ginestet et al., 2022).

476 2.4. Visuo-attentional exploration of a stimulus

477 The main goal of visuo-attentional exploration in the
 478 model is to favor efficient letter perception accumula-
 479 tion during processing. For this purpose, the model au-
 480 tomatically selects the visuo-attentional parameter val-
 481 ues that would allow gaining more information on letter
 482 identity during a given exposure. The entropy of proba-
 483 bility distributions in the letter perceptual submodel is
 484 computed to estimate the quality of perceptual repre-
 485 sentations. The entropy is close to maximal during the
 486 first iterations of processing due to limited information
 487 on letter identity within the input string. Conversely, it
 488 would be small if letters were perfectly perceived (i.e.,
 489 if perceptual representations were Dirac probability dis-
 490 tributions). Thus, a decrease in entropy characterizes
 491 letter identity information gain at the perceptual level.
 492 Measuring entropy for each letter position allows iden-
 493 tifying those letters for which perceptual information is
 494 lacking, thus indicating where attention should shift to
 495 significantly decrease entropy. How this might be per-
 496 formed with an optimization approach was described
 497 in a previous study (Ginestet et al., 2022). However,
 498 optimizing information gain entailed systematically ex-
 499 ploring the parameter space to predict the entropy de-
 500 crease for all possible combinations of visual attention
 501 parameters. This was computationally costly. Here,

502 instead, we used a heuristic-based, approximate algo- 553
503 rithm that provides visuo-attentional exploration behav- 554
504 iors that are qualitatively comparable to those produced 555
505 by our previous algorithm (a quantitative assessment of 556
506 this approximation is beyond the scope of the current 557
507 paper).

508 Here, we more specifically focus on how location of 559
509 the attentional focus moves over the input string to in- 560
510 crease the gain of information on letter identity at each 561
511 exposure and boost perceptual evidence accumulation. 562
512 We then expose how visual attention dispersion is af- 563
513 fected during processing and then provide an illustra- 564
514 tive example, through the processing of the novel word 565
515 “HOLPING”.

516 *Displacement of the visuo-attentional focus during ex-* 567
517 *ploration.* The heuristic algorithm proceeds as follows. 568
518 Initially, the position of the gaze and attentional fo- 569
519 cus μ_A^t is set according to stimulus length and atten- 570
520 tional quantity (note that gaze position always coincide 571
521 with the attentional focus position in the simulations). 572
522 Following eye movement behavioral findings (Rayner, 573
523 1998; Vitu et al., 1990), the attentional focus is located 574
524 slightly left of the word center, except for the smallest 575
525 value of attention quantity ($Q_A = 0.5$), for which the 576
526 initial position is located on the first letter of the word. 577
527 This shift towards the beginning of words was motivated 578
528 by the fact that virtually a single letter could be pro- 579
529 cessed at once in this condition, so that no information 580
530 could accumulate on the initial letter of the input stimu- 581
531 lus when the focus of attention was located farther away 582
532 on the right.

533 Then, at each time-step, the difference in entropy, 584
534 between the probability distributions of the perceptual 585
535 representation of the letter under the attentional focus 586
536 and all other positions is computed. When this differ- 587
537 ence exceeds a given threshold T_{shift} (empirically set 588
538 to 1.5 nats, with 1 nat the unit for information quan- 589
539 tity when entropy is computed using the natural loga- 590
540 rithm, as we do, instead of the more usual bit when it 591
541 is computed with the base 2 logarithm), then a visuo- 592
542 attentional shift is initiated towards that position. As 593
543 a result, except for the initial position of the focus of 594
544 attention, all subsequent displacements of the attention 595
545 focus are computed by the model depending on the 596
546 quality of identity evidence previously accumulated at 597
547 the perceptual level. As in the terminology of eye move- 598
548 ment studies, we will refer to time intervals when atten- 599
549 tion does not move as an “attentional fixation”, between 600
550 attentional displacements, and therefore count the num- 601
551 ber of attentional fixations.

552 As previously (Ginestet et al., 2022), the entropy dif-

ference was modulated by a motor cost parameter, noted 602
 α . This parameter considers the magnitude of the next 603
displacement to penalize large attentional shifts. Sev- 604
eral displacements of the focus of visual attention, thus 605
several attentional fixations, can occur in a single ex- 606
posure, as far as each displacement contributes to min- 607
imize entropy. Visuo-attentional exploration is stopped 608
whenever the average entropy on letters falls below 609
threshold T_{avg} (also empirically set to 1.5 nat), so that 610
letter identity processing is considered terminated for 611
the current exposure. 612

Modulation of visual attention dispersion during ex- 613
ploration. The model also automatically adjusts atten- 614
tional dispersion during the exploration of the input let- 615
ter string. The initial dispersion of visual attention is 616
set to its default value $\sigma_A^t = 1.75$. At the end of the 617
first displacement of the visuo-attentional focus during 618
attentional fixation, a new value is selected by the explo- 619
ration algorithm as a function of information accumula- 620
tion speed during this first attentional fixation, relative 621
to a “reference” information accumulation profile. 622

This reference profile was obtained as follows: for 623
each length, we randomly selected 100 words from the 624
lexicon, and performed letter and word recognition dur- 625
ing 1,000 iterations, with a single fixation, and all pa- 626
rameters of the model at their default values. In partic- 627
ular, gaze and attention position were slightly left of the 628
center position. We then measured the evolution of en- 629
tropy for all these words, and computed their average. 630
An example reference profile is shown Figure 3 (green 631
curve of top left plot). 632

At the end of the first attentional fixation, if informa- 633
tion accumulation was faster than in the reference, the 634
model adopts a large attentional dispersion for the rest 635
of stimulus exploration. If, on the other hand, informa- 636
tion accumulation was slower, attentional dispersion 637
is reduced, so that fewer letters are processed in each 638
attentional fixation. To compare the current entropy de- 639
crease with the reference one, their ratio is computed; 640
we have empirically defined a relation that yields atten- 641
tion dispersion for subsequent attentional fixations as a 642
function of the entropy ratio (Figure 3, top right). The 643
value of the adjusted attention dispersion parameter σ_A^t 644
is computed once at the end of the first attentional fix- 645
ation and then applied for all subsequent fixations until 646
termination. 647

In the visuo-attentional submodel, the parameters for 648
attention quantity Q_A and attention dispersion σ_A^t can 649
mathematically be manipulated independently. How- 650
ever, the visual exploration algorithm induces a strong 651
correlation between them. Indeed, as we have just de- 652

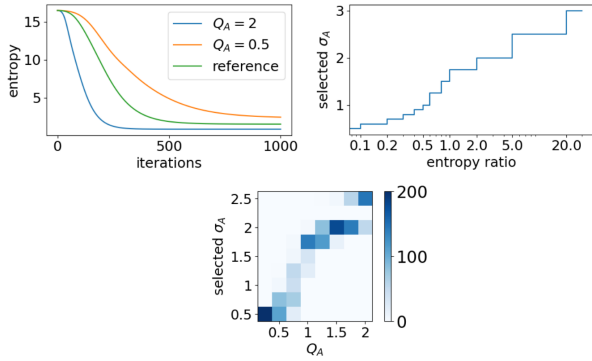


Fig. 3: Illustration of the modulation of visual attention dispersion during exploration. Top left: Evolution of the letter entropy over time. The green curve represents the reference entropy profile; the other two represent entropy evolution when the model is presented with the word “IMAGE”, for two different values of Q_A . Top right: Values of dispersion parameter σ_A^t selected by the visual exploration algorithm, as a function of the entropy gain ratio between stimulus and reference processing at the end of the first attentional fixation. Bottom: Values of dispersion parameter σ_A^t selected by the visual exploration algorithm, as a function of parameter Q_A . Color indicates how many words used each value of σ_A^t .

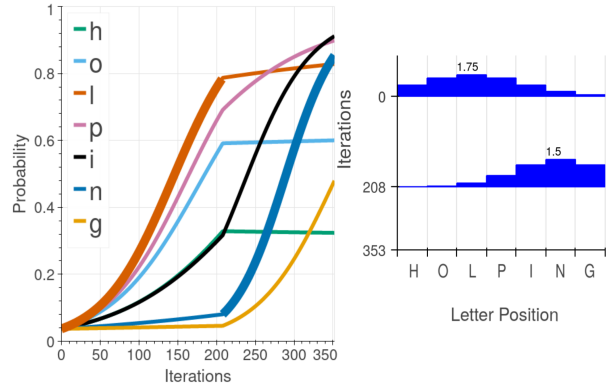


Fig. 4: Illustration of the visuo-attentional exploration algorithm on stimulus “HOLPING”. Left plot: Probability of perceived letters (y-axis) at each position, as a function of simulated time (x-axis). Each curve represents the probability value of the most likely letter hypothesis, at each position. Curves are color coded according to position (green curve for position 1, yellow curve for position 7, etc.). Curves are in thick lines when the focus of visual attention is on the position that they correspond to. Right plot: Evolution over time (y-axis) of the visuo-attentional distribution over the stimulus positions (x-axis). Letters at each positions are recalled at the bottom of the plot (“H” in position 1, etc.) Time indices indicated on the y-axis are beginnings of attentional fixations, for which the visuo-attentional distribution is the one depicted by the corresponding box plots, with its dispersion indicated by a number (e.g., between iterations 0 and 208, the focus of attention was on letter L at position 3; attention dispersion was 1.75). Box height indicates the attention allocated at each position).

604 scribed, attention dispersion σ_A^t is selected as a function
 605 of information accumulation speed, which is itself mod-
 606 ulated by attention quantity Q_A . Figure 3 (bottom plot)
 607 illustrates the correlation between the two parameters
 608 on an independent experimental dataset. This dataset
 609 was composed of 200 8-letter words that were randomly
 610 extracted from the ELP database (Balota et al., 2007).
 611 As illustrated, the smaller the visual attention quantity
 612 Q_A , the smaller the adopted attentional dispersion σ_A^t .
 613 In the rest of this paper, we consider Q_A as our variable
 614 of interest, to study its effect on the predicted behav-
 615 ior, while σ_A^t is considered as a dependent, constrained
 616 variable.

617 *Illustration: visuo-attentional exploration of the novel*
 618 *word “HOLPING”.* Figure 4 illustrates the dynamics
 619 of visuo-attentional exploration (right plot) and how let-
 620 ter identity information evolves over time at the percep-
 621 tual level (left plot), for the novel word “HOLPING”
 622 at the first exposure, with attention quantity $Q_A = 1$.
 623 At the beginning of processing (iteration 0), the dis-
 624 tribution of visual attention is characterized by a focus
 625 aligned on the third letter of the 7-letter input word and a
 626 default value dispersion $\sigma_A^t = 1.75$. During the 208 itera-
 627 tions of this first attentional fixation, letter identity in-
 628 formation gradually accumulates at the perceptual level.
 629 As can be seen on Figure 4 (left plot), during this period,
 630 identity evidence accumulates rapidly for the letter un-
 631 der the focus of attention and less so for other letters,

632 as a function of their distance from the focus of atten-
 633 tion. As a result, during the first attentional fixation,
 634 only very few letter identity information accumulates
 635 for the two final letters that are the most distant from
 636 the focus of attention.

637 At iteration 208, attention shifts to position 6 (i.e., on
 638 the letter N of “HOLPING”), a position that simulta-
 639 neously maximizes the expected entropy gain and min-
 640 imizes the motor cost associated with visual attention
 641 displacement. Given that identity evidence accumu-
 642 lated relatively efficiently for most letters during the first
 643 attentional fixation, visual attention dispersion is only
 644 slightly adjusted, leading to a σ_A^t value of 1.5. As can be
 645 seen on Figure 4 (left plot), the consequence of a visual
 646 attention shift at iteration 208 is twofold. First, identifi-
 647 cation of the letters at and immediately around the new
 648 attentional focus is boosted, yielding a sharp increase in
 649 identification probability for the final letters (“ING”);
 650 second, identification probability begins to decrease for
 651 the initial letters that no longer receive attention. At the
 652 end of the second attentional fixation (iteration 353), the
 653 termination criterion based on threshold T_{avg} is met, so
 654 that visual exploration and processing of the stimulus
 655 end.

656 At the end of processing, lexical membership evaluation
657 tion assessed the stimulus word as being a novel word,
658 so that a new lexical representation was created. This
659 lexical representation corresponds to knowledge accu-
660 mulated on letter identity during processing. For the
661 novel word “HOLPING”, the new memory trace will be
662 relatively complete, providing some identity informa-
663 tion on all the letters of the input string. However, none
664 of the input letters were perfectly identified at the first
665 exposure (none reached Dirac probability) and some let-
666 ters were better identified than others, thus leading the
667 possibility to improve lexical knowledge for this item
668 during subsequent exposures. To evaluate simulations,
669 two measures characterizing processing at the first ex-
670 posure are considered: a measure of processing time (in
671 this example with the novel word “HOLPING”, 353 it-
672 erations) and a measure of the number of attentional fix-
673 ations during this processing time (here, 2).

674 3. Method

675 3.1. Material

676 Seven hundred words were selected from the model’s
677 lexical database to serve as stimuli for the current study.
678 The words varied in length from 4 to 10 letters. We used
679 the Gurobi problem solver (Gurobi Optimization LLC,
680 Beaverton, Oregon, USA; Gurobi Optimization, LLC
681 2021), to select one hundred words, for each length,
682 so that they were matched in frequency and belonged
683 to the Noun grammatical category. The selected words
684 were of medium frequency, varying between 3.6 and 3.7
685 occurrences per million words (the average frequency
686 of the whole lexicon was 3.63 occurrences per million
687 words). To exclude any potential additional effect of
688 neighborhood, all target word neighbors (i.e., all the
689 words that differed from target words by a single let-
690 ter) were excluded from the lexicon, thus resulting in a
691 set of stimuli without orthographic neighbors. This re-
692 moved 1,983 words from the 79,673 (2.5%) words of
693 the lexicon. Removing the orthographic neighbors al-
694 lowed studying the length effect while excluding con-
695 founding factors. Indeed, short words typically have
696 many more orthographic neighbors than long words, so
697 that the number of neighbors cannot be equated for sets
698 of words that strongly differ in length.

699 For the current experiment, this set of 700 words was
700 used twice. They were considered once as known words
701 – thus belonging to the model’s lexical word knowl-
702 edge – and once as novel words, in which case they
703 were removed from the model’s lexical database. This
704 was done to ensure a perfect matching between the char-
705 acteristics of stimuli, independently of their status as

706 known words or novel words; this also ensures that
707 stimuli considered as novel words are realistic, in the
708 sense that, for instance, they are orthographically legal.
709 The list of stimuli can be found in Appendix A.

710 3.2. Procedure

711 The model was used to simulate the visuo-attentional
712 exploration of the 700 stimuli, twice each, as each was
713 once considered a known word and once as a novel
714 word, for a total of 1,400 simulations. This was re-
715 peated for seven possible values of attention quantity
716 Q_A (0.5, 0.75, 1, 1.25, 1.5, 1.75 and 2). In each sim-
717 ulation, the same stimulus was presented five times to
718 the model: at each of these exposures, we simulated the
719 visual-attentional exploration of the stimulus, and the
720 subsequent updating of an existing orthographic trace,
721 or the creation of a new one.

722 From each simulated exposure, we measured two
723 variables of interest. First, a measure of Processing
724 Time (PT) was computed as the number of iterations
725 occurring before the termination criterion was met. Sec-
726 ond, we measured the Number of Attentional Fixations
727 (NAF) performed by the model in the same time inter-
728 val. The length effect was quantified by the slope be-
729 tween performance on the two measures of interest for
730 the shortest and the longest items, item length being es-
731 timated in number of letters (4 versus 10 letters).

732 3.3. Statistical analyses

733 The simulated Processing Times were analyzed us-
734 ing generalized linear models (*glm* function; R Core
735 Team 2020) with a Gamma family and an inverse link.
736 To select the most appropriate link function, we tested
737 several possibilities (“identity”, “inverse” and “log”)
738 and analyzed the results of the subsequent models:
739 we chose the model that minimized both the result-
740 ing AIC (Akaike Information Criterion; Akaike, 1973)
741 and the Fisher Scoring (number of iterations required
742 for the model to converge). To analyze the NAF, we
743 followed the suggestion of Harris et al. (2012) and
744 used a generalized Poisson regression (*vglm* function;
745 R Core Team 2020), as the data were underdispersed
746 (dispersiontest function; R Core Team 2020). All
747 statistical models and simulated results are provided as
748 Supplementary Material¹

749 First, we used two models to compare PT and NAF
750 for words and novel words at the first exposure, in which
751 Attention Quantity (7 Q_A values), Item Type (novel

¹Open access availability for Supplementary Material files:
<https://osf.io/g8cbf/>.

word vs. known word) and Item Length (from 4 to 10 letters) were included as fixed factors. For the sake of clarity, results are first presented while focusing on the lexicality effect, then, on the length effect.

Second, we used two models to analyze PT and NAF across exposures, but for the novel words only, with Attention Quantity (Q_A), Item Length and Exposure Number (from 1 to 5) as fixed factors. The results are first presented while focusing on the interaction between Q_A and the number of exposures, in which case PT and NAF are expressed per letter, then focusing on the length effect for the two variables of interest (PT and NAF).

4. Simulation results

For the known words, the process of orthographic learning was always successful, for all Item Lengths and Attention Quantity Q_A values. For novel words, orthographic learning sometimes failed. This occurred when a novel word was erroneously categorized as a known word, so that the orthographic trace of the most activated known word (typically an orthographically similar word) was updated. Erroneous learning further occurred when a previously encountered novel word was once more categorized as novel during a subsequent exposure, so that a new, extraneous trace was created and the orthographic trace previously created for this same novel word was not updated.

The success rates for novel word learning are provided in Table 1 for the different Q_A values and lengths. While all the shorter novel words (from 4 to 6 letters) were successfully learned regardless of Q_A , learning errors were observed for longer items. As shown in Table 1, the success learning rate increased as the Attention Quantity Q_A increased. For each Q_A value, stimuli that generated learning errors were excluded from all further analyses.

The effect of Q_A on stimuli processing is described in the next two sections. We first focus on processing at the first exposure to describe how Attention Quantity affects PT and the NAF depending on Item Type (novel words vs. known words) and Item Length (from 4 to 10 letters). Given the high level of performance of the model for known words from the first exposure, in the second section, we focus on novel word processing alone to describe how the Item Length effect evolves across the five exposures depending on Attention Quantity. Note that all the results reported in the following sections were derived from the same data set using a single statistical model for each measure. They are presented in different sections for the sake of clarity.

4.1. Processing of known words and novel words at the first exposure

The effect of Q_A on PT and NAF for the two types of items at the first exposure is illustrated in Figure 5. Keep in mind that stimuli are of variable length, and thus induce very different PT and NAF. For the coherence of the figure, and since we are not focusing on the length effect for now, both PT and NAF were normalized by word length. Novel words were processed slower than known words ($\beta = -5.6e-4, t = -14.70, p < .001$). Regardless of Item Type, average PT decreased when Attention Quantity increased ($\beta = 6.5e-4, t = 59.60, p < .001$), varying from 188 iterations per letter on average for $Q_A = 0.5$ to 59 iterations per letter on average for $Q_A = 2$. More importantly, the Attention Quantity (Q_A) by Item Type interaction was significant ($\beta = -1.1e-4, t = -8.80, p < .001$), showing that PT decreased more for novel words than for known words as the Attention Quantity increased. Average PT varied from 261 iterations per letter for $Q_A = 0.5$ to 70 iterations per letter for $Q_A = 2$ for the novel words and from 127 to 47 iterations per letter for the known words. As a result, the difference in PT between known words and novel words, that is the lexicality effect on PT, decreased when more attention quantity was available for processing.

Similar effects characterized NAF performance. The Attention Quantity (Q_A) by Item Type interaction was significant ($\beta = -0.078, z = -3.24, p = .001$). Post-hoc analysis showed that Attention Quantity (Q_A) affected NAF for the novel words ($\beta = -0.095, z = -6.47, p < .001$) but not for the known words ($\beta = -0.016, z = -0.86, p = .392$). With respect to novel words, average NAF varied from 1.17 NAF per letter for $Q_A = 0.5$ to 0.44 NAF per letter for $Q_A = 2$. With respect to known words, average NAF varied from 0.52 NAF per letter for $Q_A = 0.5$ to 0.33 NAF per letter for $Q_A = 2$. Thus, the lexicality effect on NAF was modulated by Attention Quantity, so that the difference in NAF between known and novel words decreased when Attention Quantity (Q_A) increased. Otherwise, the main Item Type effect was significant; more attentional fixations were observed on novel words than on known words ($\beta = 0.28, z = 2.87, p = .004$).

At the first exposure, the effect of Q_A on PT and NAF for the two types of items depending on Item Length is illustrated in Figure 6. This figure illustrates the same data as the previous one, and corresponds to the same statistical analyses. However, the graphical representation here focuses on the impact of Item Length on the two measures of PT and NAF. With respect to PT, the Item Length effect was modulated by Attention

Table 1: Successful learning rate, in the learning simulation, for novel words (successful learning rate is 1.0 for words).

Q_A	Length							
	4L	5L	6L	7L	8L	9L	10L	
0.5	1.0	1.0	1.0	0.88	0.81	0.68	0.56	
0.75	1.0	1.0	1.0	0.97	0.95	0.80	0.73	
1	1.0	1.0	1.0	0.96	0.96	0.83	0.71	
1.25	1.0	1.0	1.0	0.97	0.97	0.85	0.80	
1.5	1.0	1.0	1.0	0.97	0.99	0.85	0.85	
1.75	1.0	1.0	1.0	0.97	0.99	0.91	0.88	
2	1.0	1.0	1.0	0.98	0.98	0.93	0.89	

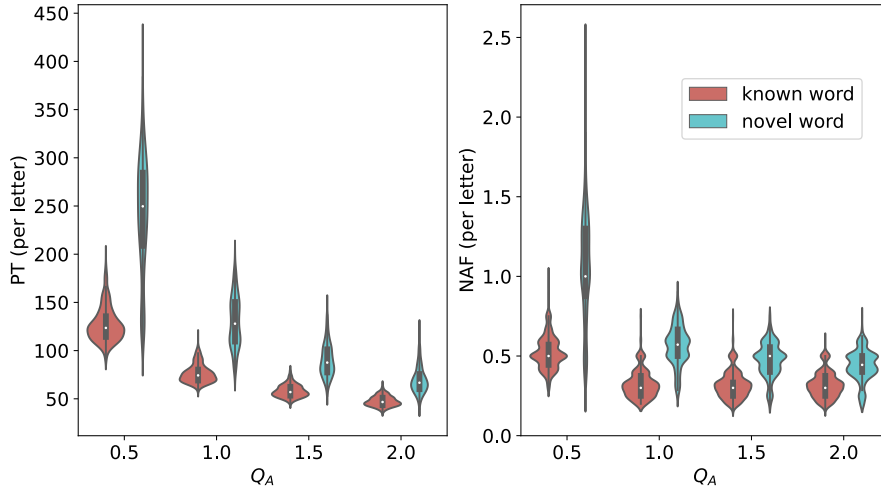


Fig. 5: Processing Time (PT, left) and Number of Attentional Fixations (NAF, right) per letter (y-axes), depending of Item Type (known words, in light blue, or novel words, in dark blue), as a function of visual Attention Quantity (Q_A values, x-axes). For each measure, a “violin plot” depicts the distribution of obtained values, with wider portions indicating higher density of values. The central dot represents the median of the distribution of values.

854 Quantity (Q_A): it was larger when Attention Quantity
855 was smaller ($\beta = -4.5e-5, t = -33.91, p < .001$).
856 There was no larger Item Length effect on PTs for the
857 novel words than for the known words, as shown by the
858 non significant Item Type by Item Length interaction
859 ($\beta = -2.8e-6, t = -0.61, p = .545$). This is due to
860 the range of explored Q_A values, in which large values
861 yield a floor effect on Processing Times; the interaction
862 is significant when considering only small Q_A values
863 (e.g., when $Q_A < 1$). However, the Attention Quan-
864 tity by Item Type by Item Length double interaction
865 was significant ($\beta = 6.92e-6, t = -4.55, p < .001$),
866 showing that the Length effect on PT was larger for
867 novel words than for words when Attention Quantity
868 (Q_A) was smaller. Otherwise, the main Item Length ef-
869 fect on PTs was significant (varying from 431 iterations
870 for 4-letter items to 975 iterations for 10-letter items;
871 $\beta = -7.0e-5, t = -16.81, p < .001$).

872 As shown on Figure 6, the Length effect on NAF
873 was greater for novel words than for known words
874 ($\beta = 0.093, z = 7.42, p < .001$), and greater for the
875 lower values of Attention Quantity ($\beta = -6.6e-3, z =$
876 $-2.65, p = .008$). However, neither the Attention Quan-
877 tity by Length interaction nor the Attention Quantity by
878 Length by Item Type double interaction were significant
879 ($\beta = -4.3e-3, z = -1.37, p < .170$). The main effect of
880 Length was significant ($\beta = 0.10, z = 10.43, p < .001$),
881 varying from 2.18 NAF for 4-letter items to 4.52 for 10-
882 letter items.

4.2. Evolution of the processing of novel words across exposures

883 Figure 7 illustrates the effect of both Q_A and the Num-
884 ber of Exposures on novel words’ PT and NAF. As
885 shown on Figure 7 (left), PT decreased across Expo-
886 sures ($\beta = 9.5e-5, t = 11.0, p < .001$), varying from
887
888

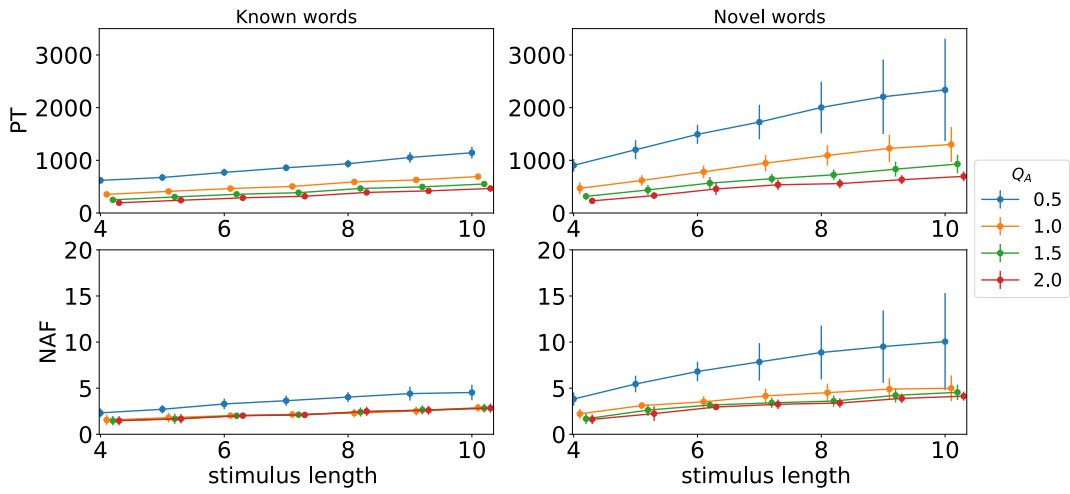


Fig. 6: Measures of visuo-attentional exploration (PT, top row and NAF, bottom row, on y-axes), at the first exposure, for known (left column) and novel words (right column), as a function of stimulus length (x-axes) and Attention Quantity Q_A (colored curves, from blue ($Q_A = 0.5$) to pink ($Q_A = 2.0$)). Error bars represent the data's standard deviation. The curves are slightly shifted horizontally from each other to ensure that the error bars are readable in the presence of overlap.

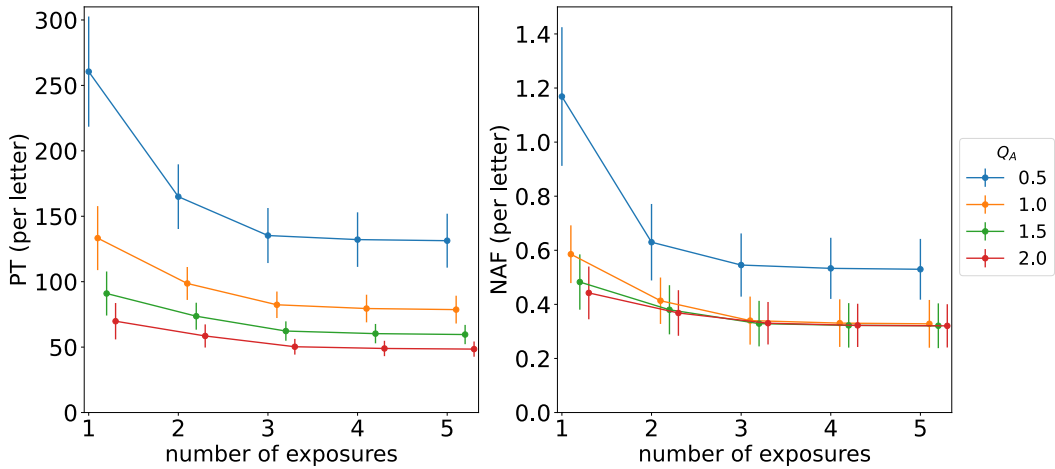


Fig. 7: Measures of visuo-attentional exploration (left, PT, in number of iterations per letter; right, NAF, in number of attentional fixations per letter; on y-axes) across exposures (x-axes) for novel words. Error bars represent the data's standard deviation. Curves are slightly shifted horizontally from each other to ensure that the error bars are readable in the presence of overlap. Each curve refers to a given visual Attention Quantity (Q_A), from 0.5 (blue) to 2.0 (pink).

889 128 iterations per letter on average at the first exposure
 890 to 74 iterations per letter at the fifth exposure. The
 891 Attention Quantity (Q_A) by Exposure interaction was sig-
 892 nificant ($\beta = 4.6e-5, t = 17.10, p < .001$), showing
 893 that the decrease in PT across exposures was stronger
 894 when visual Attention Quantity (Q_A) was more limited.
 895 Processing Times varied from 261 iterations per letter
 896 to 130 iterations per letter across the five exposures for

897 $Q_A = 0.5$, from 70 iterations per letter to 48 iterations
 898 per letter for $Q_A = 2$. For all Q_A values, Processing
 899 Time stabilized after a few exposures, but the PT value
 900 at stabilization was higher for the lower values of Q_A ,
 901 suggesting less efficient orthographic learning when the
 902 visuo-attentional quantity allocated to processing was
 903 more limited. For the lower Q_A values ($Q_A < 1$), PT
 904 after five exposures remained higher than PT at the first

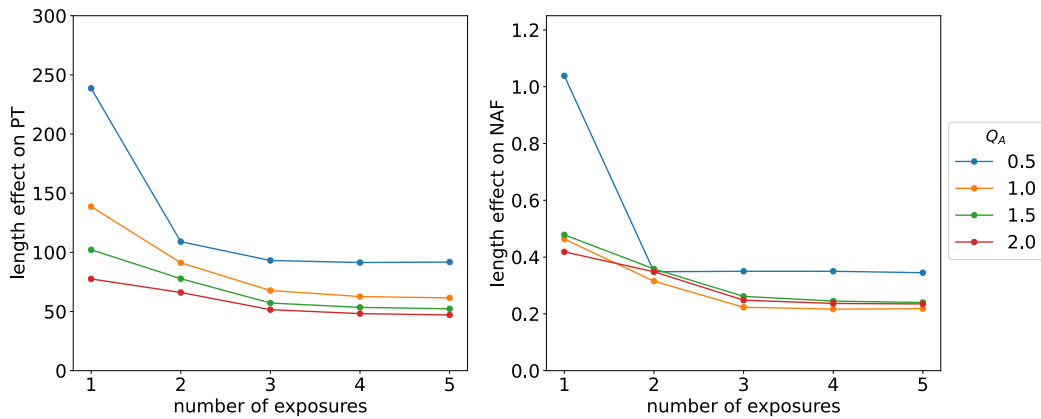


Fig. 8: Evolution of the length effect on PT (left, in number of additional iterations per additional letter) and NAF (right, on number of additional attentional fixations per additional letter), on y-axes, as a function of exposures (x-axes). Each curve refers to a given visual Attention Quantity (Q_A), from 0.5 (blue) to 2.0 (pink).

905 exposure for the higher Q_A values.

906 Different patterns characterized NAF performance. 936
 907 As shown on Figure 7 (right), neither the main effect of 937
 908 Exposure nor the Attention Quantity (Q_A) by Exposure 938
 909 interaction were significant ($\beta = -0.031, z = -1.37, p =$ 939
 910 $.172$ and $\beta = 1.4e-3, z = 0.25, p = .801$ respectively). 940

911 The plots on Figure 8 illustrate the evolution of length 942
 912 effects on novel words' PT and NAF across Exposures 943
 913 depending on Attention Quantity. As shown on Fig- 944
 914 ure 8 (left), the Exposure by Length interaction was 945
 915 significant ($\beta = 1.1e-5, t = 9.76, p < .001$), show- 946
 916 ing that the difference in PT between the shortest and 947
 917 the longest words was reduced across exposures. This 948
 918 reduction was further modulated by visual Attention 949
 919 Quantity (Q_A), as shown by the significant Attention 950
 920 Quantity by Length by Exposure double interaction ($\beta =$ 951
 921 $-4.3e-6, t = -13.11, p < .001$). The length effect on 952
 922 PTs diminishes faster across exposures when Attention 953
 923 Quantity was lower. 954

924 The same pattern was observed regarding NAF (see 956
 925 Figure 8, right). Both the Exposure by Length inter- 957
 926 action ($\beta = -0.031, z = -10.10, p < .001$) and the At- 958
 927 tention Quantity by Exposure by Length double interac- 959
 928 tion ($\beta = 3.6e-3, z = 4.95, p < .001$) were significant. 960
 929 The NAF was far more important for the longest than 961
 930 the shortest words at the first (6.24 vs. 2.23 for the 10- 962
 931 and 4-letter words respectively) than at the fifth expo- 963
 932 sure (3.08 vs. 1.72) and the NAF difference between the 964
 933 longest and the shortest words decreased faster across 965
 934 Exposures when (Q_A) was lower. 966

935 5. Discussion

In the present paper, computational modeling was used to examine the role of visual attention in the transition from more serial to more parallel letter-string processing. We used the BRAID-Learn model, a model of orthographic processing that includes word recognition and orthographic learning mechanisms, as an experimental substitute.

Simulations showed that lexicality and length effects on PT and NAF decreased when larger visual attention quantity was available for processing. Orthographic learning was less successful when visual attention quantity was smaller and the input novel word longer. The evolution patterns of orthographic processing across exposures were also affected by visual attention quantity. Repeated exposure to the same novel word resulted in a larger decrease of PT and NAF when the quantity of visual attention was smaller. In the same way, smaller visual attention quantity yielded a larger decrease of the length effect on PT and NAF with repeated exposure to the same novel word. Overall, the model predicts that variations in visual attention quantity would significantly affect letter string processing and orthographic learning.

The advantage of computational modeling is to offer the opportunity to examine the effect of a single parameter manipulation, here visual attention quantity Q_A , on orthographic processing while controlling for all the other effects, either inherent to the system (like visual acuity or lateral interference) or to the input stimuli (like frequency or lexical neighborhood). However, isolating a single mechanism in this manner is easier in a compu-

967 tational model than in behavioral studies. Furthermore, 1017
968 the amount of visual attention available for processing 1018
969 is not easy to measure in humans, even though estimat- 1019
970 ing it in reference to the Theory of Visual Attention has 1020
971 been attempted (Bogon et al., 2014; Bundesen, 1990). 1021

972 Therefore, to evaluate the plausibility and relevance 1022
973 of the model’s predictions, we will concentrate on the 1023
974 orthographic processing mechanisms that are responsi- 1024
975 ble for the simulated lexicality and length effects, first 1025
976 without considering the effect of Q_A variations. Second, 1026
977 provided a close relationship between the model’s gener- 1027
978 al predictions and behavioral findings, we will discuss 1028
979 to what extent the evolution of the lexicality and length 1029
980 effects on PT and NAF depending on visual attention 1030
981 quantity provides insights on the serial-to-more-parallel 1031
982 transition and is compatible with available behavioral 1032
983 evidence. 1033

984 *5.1. Lexicality and word length effects irrespective of* 1035 985 *Q_A* 1036

986 We focused on the two effects of lexicality and word 1038
987 length, as markers of serial processing. The lexicality 1039
988 effect in the model directly follows from top-down influ- 1040
989 ence of word knowledge that speeds up letter identifica- 1041
990 tion at the perceptual level and facilitates processing for 1042
991 the input letter strings that match an orthographic repre- 1043
992 sentation. The length effect in the model follows from 1044
993 the fact that the same amount of visual attention spreads 1045
994 over the input letter string whatever its length, so that 1046
995 less attention is allocated to each letter in longer stim- 1047
996 uli. As a result, letter identity information accumulates 1048
997 less efficiently at the perceptual level for longer than 1049
998 for shorter stimuli, which increases PT and NAF during 1050
999 visuo-attentional exploration of the input string. How- 1051
1000 ever, partial identity information accumulated at the per- 1052
1001 ceptual level through visuo-attentional exploration can 1053
1002 be compensated by top-down lexical information, so 1054
1003 that known words suffer lesser length effects than novel 1055
1004 words, that have no orthographic representation (at the 1056
1005 first exposure). These simulated length and lexicality 1057
1006 effects, and their interaction, are coherent with many be- 1058
1007 havioral findings from studies on eye movements, word 1059
1008 recognition and reading (Barton et al., 2014). In par- 1060
1009 ticular, longer fixation duration and a higher number of 1061
1010 fixations are reported in longer than shorter words (Hau- 1062
1011 tala et al., 2011; Joseph et al., 2009; Kliegl et al., 2004; 1063
1012 Loberg et al., 2019; McDonald, 2006; Rayner, 1998). 1064
1013 Readers spend more times fixating novel words (Chaf- 1065
1014 fin et al., 2001; Williams and Morris, 2004) and show a 1066
1015 larger length effect on these items than on known words 1067
1016 (Lowell et al., 2014).

In the same way, some general learning effects like the reduction of PT and NAF with repeated exposure to novel words (independently of Q_A) directly follow from the combined effects of visuo-attentional exploration and lexical feedback. At the first exposure, perceptual information on letters is only based on stimulus sensory processing, since no lexical representation is available yet for this word. From the second exposure, perceptual information benefits from the influence of the newly created orthographic representation. Improvement of the novel word orthographic representation across exposures results in an increase of lexical feedback that enhances letter identification. As a result, orthographic learning in the model is characterized by a decrease in PT and NAF, which is consistent with behavioral findings from studies on the evolution of eye movement patterns in conditions of orthographic learning (Ginestet et al., 2020; Joseph and Nation, 2018; Joseph et al., 2014; Pagan and Nation, 2019; Pellicer-Sanchez, 2016).

In our simulations, we further observed a decrease in the length effect with repeated exposure to the same novel word. This follows from the fact that better-specified orthographic representations have higher influence on letter perceptual information and that lexical feedback is particularly critical when bottom-up perceptual identity information accumulates slowly, which more likely occurs for longer than shorter words. Obviously, when the attentional fixation is directed towards initial letters, final letters do receive less attention in longer than in shorter words. As a direct consequence, perceptual information accumulates more slowly for longer words that are thus more dependent on lexical feedback. Several behavioral studies have reported a reduction of the length effect on reading latency after a few repeated exposures to novel words (Kwok and Ellis, 2014; Maloney et al., 2014; Suárez-Coalla et al., 2014). Behavioral evidence that longer words progressively tended to be read as quickly as shorter words was interpreted as a marker of orthographic learning, suggesting that more and more letters within the input string were simultaneously processed.

5.2. *Modulation of lexicality and length effects by attention quantity*

Our main contribution in the present paper was to evaluate the influence of visual attention quantity on orthographic processing. The model predicts that the two lexicality and length effects are modulated by visual attention quantity, thus suggesting that the total amount of visual attention available for processing further contributes to the serial-to-more-parallel processing transi-

tion. In the model, the amount of visual attention quantity deployed for processing at the first attentional fixation modulates the speed of letter identity perceptual identification and the number of letters that fall under the deployed attention. At the second fixation, visuo-attentional dispersion is modulated according to previous information accumulation speed. Fast accumulation of identity information for the higher Q_A values leads to adopt larger visual attention dispersion. A higher number of letters are then simultaneously identified at each new fixation, leading to more parallel processing. To the contrary, attentional dispersion is narrowed when identity information accumulated laboriously at the first attentional fixation. Then, only a few letters can be successfully identified at each subsequent fixation, leading to more serial processing.

Although it is difficult to directly measure the visual attention quantity in humans, the impact of perceptual processing speed and multi-letter parallel processing on behavioral performance have been investigated by reference to two theoretical frameworks, namely the Theory of Visual Attention (Bundesen, 1990; Bundesen and Habekost, 2014) and that of visual attention span (Bosse et al., 2007; Valdois, 2022; Valdois et al., 2004). Moreover, behavioral studies have established a link between perceptual processing speed and VAS, suggesting that lower VAS performance related to slower perceptual processing (Bogon et al., 2014; Dubois et al., 2010; Ginestet et al., 2020; Lobier et al., 2013). The plausibility of the model's predictions with respect to variations in visual attention quantity can therefore be questioned in the light of available behavioral evidence on how perceptual processing speed and VAS affect letter-string processing and orthographic learning.

The model predicts that individuals with smaller visual attention quantity would be more prone to rely on serial processing, thus showing higher lexicality and length effects on processing time and number of fixations while reading. The studies carried out by reference to the Theory of Visual Attention (Bundesen, 1990; Bundesen and Habekost, 2014) provide some support to this prediction. Perceptual processing speed was consistently found reduced in brain-damaged individuals showing excessive reliance on serial processing (Habekost, 2015). In particular, perceptual processing speed is markedly reduced in letter-by-letter readers who otherwise exhibit exaggerated word length effects on naming and lexical decision latencies, and eye movement measures (Barton et al., 2014; Behrmann et al., 2001). However, we lack direct evidence that word processing and the oculomotor pattern in letter-by-letter readers are related to their perceptual processing speed

(or VAS). Future studies should more directly evaluate whether differences in perceptual processing speed would predict the amplitude of the length effect in letter-by-letter readers.

Lower visual attention quantity might further account for stronger reliance on serial processing in developmental dyslexia. Several studies suggest that individuals with developmental dyslexia exhibit a reduction in perceptual processing speed (Habekost, 2015; Stefanac et al., 2019; Stenneken et al., 2011) and in visual attention span (Bosse et al., 2007; Germano et al., 2014; Zoubrinetzky et al., 2014). Furthermore, it is well documented that a larger word-length effect on naming, lexical decision and oculomotor measures is a consistent finding in developmental dyslexia (De Luca et al., 2002; Martens and de Jong, 2008; Spinelli et al., 2005; Zoccolotti et al., 2005). However, once again, direct evidence that reduced processing speed or VAS affects the lexicality or length effects in developmental dyslexia is scarce. An exaggerated length effect has been described in association with reduced VAS in some case studies of developmental dyslexia (Valdois et al., 2011, 2003) and a group study has shown that the number of fixations (but not fixation duration) in text reading increased in dyslexic individuals with lower VAS (Prado et al., 2007). A more rigorous assessment of the model predictions would require to systematically evaluate whether a VAS or perceptual processing speed deficit in developmental dyslexia is associated to excessive length and lexicality effects.

However, the main prediction of the model is that differences in visual attention quantity should affect the transition from serial-to-more-parallel processing. Relevant behavioral evidence would then come from changes in reading patterns across grades and from orthographic learning studies. Only piecemeal behavioral information can be related to the model's prediction. There is evidence that VAS abilities increase across grades (van den Boer et al., 2015; van den Boer and de Jong, 2018; Bosse and Valdois, 2009; Huang et al., 2019). The large decline in word-length effect observed in typical readers as they learn to read might thus suggest a decrease in word length effect with growth in VAS skills. Unfortunately, we lack direct behavioral evidence for such a relationship across grades. However, van den Boer et al. (2013) showed that variations in VAS skills in second grade children predicted variations in length effect on their reading latencies. This finding and the consistently reported relationship between VAS and reading fluency (van den Boer and de Jong, 2018; Bosse and Valdois, 2009; Chan and Yeung, 2020; Chen et al., 2016; Lobier et al., 2013; Valdois et al., 2021b, 2019;

1172 Zhao et al., 2018) suggest that VAS would contribute to 1223
1173 the degree of reliance on serial processing. 1224

1174 To our knowledge, no study investigated the relation- 1225
1175 ship between VAS (or processing speed) and the lex- 1226
1176 icality effect. Antzaka et al. (2017) examined skilled 1227
1177 readers' pseudo-word reading in conditions of very brief 1228
1178 presentation duration that prevented serial processing. 1229
1179 They showed that the adult readers who played action 1230
1180 video games and had larger VAS than non-players could 1231
1181 successfully read more pseudo-words through parallel 1232
1182 processing. As the two groups of players and non- 1233
1183 players were matched on text reading fluency, their find- 1234
1184 ings might suggest that larger VAS is associated to a 1235
1185 lower lexicality effect on processing times. Behav- 1236
1186 ioral studies on orthographic learning should be par- 1237
1187 ticularly relevant to evaluate the link between visuo- 1238
1188 attentional resources and the shift from serial-to-more- 1239
1189 parallel processing. Unfortunately, although available 1240
1190 findings convincingly show incremental orthographic 1241
1191 knowledge growth across repeated exposure to the same 1242
1192 novel word (Joseph and Nation, 2018; Joseph et al., 1243
1193 2014; Pagan and Nation, 2019; Pellicer-Sanchez, 2016), 1244
1194 neither VAS nor perceptual processing speed were si- 1245
1195 multaneously measured. A single study provided some 1246
1196 evidence of better orthographic learning skills in the 1247
1197 group of participants with higher VAS (Ginestet et al., 1248
1198 2020). 1249

1199 5.3. Conclusion and perspectives 1250

1200 The main contribution of the present modeling study 1252
1201 is twofold. First, the model provides a sophisticated 1253
1202 description of the dynamics of visuo-attentional ex- 1254
1203 ploration during printed word processing. Second, it 1255
1204 shows how the interaction of visuo-attentional explo- 1256
1205 ration and lexical knowledge contributes to the grad- 1257
1206 ual strengthening of item-specific orthographic repre-
1207 sentations as learning progresses. Decrease of the
1208 lexicality and length effect across exposures suggests 1258
1209 that the model captures some aspects of the transition
1210 from serial to more parallel processing. However, or- 1259
1211 thographic learning in the model is performed in the 1260
1212 absence of any phonological processing. This drastically 1261
1213 differs from previous modeling of orthographic learn- 1262
1214 ing through self-teaching (Pritchard et al., 2018; Ziegler 1263
1215 et al., 2014), in which successful phonological process- 1264
1216 ing was critical to acquire new orthographic knowledge
1217 and explain the transition from serial to more parallel
1218 processing.

1219 In this respect, BRAID-Learn more directly relates
1220 to the model of automaticity in reading proposed by
1221 LaBerge and Samuels (1974). LaBerge and Samuels
1222 (1974) emphasized the role of visual attention in the

processing and memorization of increasingly large or-
thographic units during the course of learning to read.
In the same way, in BRAID-Learn, the amount of vi-
sual attention quantity influences the size (in letter num-
ber) of the processed units (from individual letters to
the whole word letter-string), so that the smaller the at-
tention quantity, the smaller the number of letters pro-
cessed as a whole. However, in the absence of imple-
mented phonological component, the predictive power
of BRAID-Learn is limited. Addition of a phonologi-
cal module in BRAID-Learn, or the addition of visuo-
attentional processes in dual-route self-teaching models
(Pritchard et al., 2018; Ziegler et al., 2014), would al-
low improving the models' predictions and examining
the combined effects of visual attention and phonologi-
cal processing on both orthographic learning and the
transition from serial-to-more-parallel processing.

One could further question the relevance of our study,
in which the BRAID-Learn model was equipped with
an expert orthographic lexicon and tasked to learn a sin-
gle novel word, to provide insights on reading acquisi-
tion. Indeed, during reading acquisition, it is unclear
how the current state of the growing lexicon affects the
learning of a currently encountered novel word. We
surmise that our observations would generalize to this
situation, since, at the first encounter, top-down lexi-
cal feedback is suppressed in the BRAID-Learn model,
so that the current state of the lexicon does not affect
perceptual processing and visuo-attentional exploration.
However, the interaction with phonological processing,
would certainly matter. Current work concerns extend-
ing BRAID-Learn in this direction, to study its capac-
ity to gradually build up rich lexical knowledge, while
starting from only minimal knowledge on word-specific
orthographic representations.

Acknowledgments

This work was supported by a French Ministry of Re-
search (MESR) Ph.D. grant to AS. This work was also
supported by the French government as part of the e-
FRAN "FLUENCE" project (SV as PI) funded by the
PIA2 "Investissement d'Avenir" program handled by
the "Caisse des Dépôts et Consignations".

References

Akaike, H., 1973. Maximum likelihood identification of Gaussian
autoregressive moving average models. *Biometrika* 60, 255–265.
doi:10.1093/biomet/60.2.255. publisher: Oxford University
Press.

- Antzaka, A., Lallier, M., Meyer, S., Diard, J., Carreiras, M., Valdois, S., 2017. Enhancing reading performance through action video games: the role of visual attention span. *Scientific Reports* 7, 14563. doi:10.1038/s41598-017-15119-9. number: 1 Publisher: Nature Publishing Group.
- Balota, D.A., Yap, M.J., Hutchison, K.A., Cortese, M.J., Kessler, B., Loftis, B., Neely, J.H., Nelson, D.L., Simpson, G.B., Treiman, R., 2007. The English Lexicon Project. *Behavior Research Methods* 39, 445–459. doi:10.3758/BF03193014.
- Barton, J.J.S., Hanif, H.M., Björnström, L.E., Hills, C., 2014. The word-length effect in reading: A review. *Cognitive Neuropsychology* 31, 378–412. doi:10.1080/02643294.2014.895314. publisher: Routledge _eprint: <https://doi.org/10.1080/02643294.2014.895314>.
- Behrmann, M., Shomstein, S., Black, S.E., Barton, J.J.S., 2001. The eye movements of pure alexic patients during reading and non-reading tasks. *Neuropsychologia* 39, 983–1002. doi:10.1016/S0028-3932(01)00021-5.
- van den Boer, M., van Bergen, E., de Jong, P.F., 2015. The specific relation of visual attention span with reading and spelling in Dutch. *Learning and Individual Differences* 39, 141–149. doi:10.1016/j.lindif.2015.03.017.
- van den Boer, M., de Jong, P.F., 2018. Stability of Visual Attention Span Performance and Its Relation With Reading Over Time. *Scientific Studies of Reading* 22, 434–441. doi:10.1080/10888438.2018.1472266.
- van den Boer, M., de Jong, P.F., Haentjens-van Meeteren, M.M., 2013. Modeling the length effect: Specifying the relation with visual and phonological correlates of reading. *Scientific Studies of Reading* 17, 243–256. doi:10.1080/10888438.2012.683222. place: United Kingdom Publisher: Taylor & Francis.
- Bogon, J., Finke, K., Schulte-Körne, G., Müller, H.J., Schneider, W.X., Stenken, P., 2014. Parameter-based assessment of disturbed and intact components of visual attention in children with developmental dyslexia. *Developmental Science* 17, 697–713. doi:10.1111/desc.12150. _eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1111/desc.12150>.
- Bosse, M.L., Chaves, N., Largy, P., Valdois, S., 2015. Orthographic learning during reading: the role of whole-word visual processing. *Journal of Research in Reading* 38, 141–158. doi:10.1111/j.1467-9817.2012.01551.x.
- Bosse, M.L., Tainturier, M.J., Valdois, S., 2007. Developmental dyslexia: The visual attention span deficit hypothesis. *Cognition* 104, 198–230. doi:10.1016/j.cognition.2006.05.009.
- Bosse, M.L., Valdois, S., 2009. Influence of the visual attention span on child reading performance: a cross-sectional study. *Journal of Research in Reading* 32, 230–253. doi:10.1111/j.1467-9817.2008.01387.x.
- Bowers, J.S., Davis, C.J., Hanley, D.A., 2005. Interfering neighbours: The impact of novel word learning on the identification of visually similar words. *Cognition* 97, B45–B54. doi:10.1016/j.cognition.2005.02.002.
- Bowey, J.A., Muller, D., 2005. Phonological recoding and rapid orthographic learning in third-graders' silent reading: A critical test of the self-teaching hypothesis. *Journal of Experimental Child Psychology* 92, 203–219. doi:10.1016/j.jecp.2005.06.005.
- Bundesen, C., 1990. Theory of visual attention. *Psychological review* 97, 523–47. doi:10.1037/0033-295X.97.4.523.
- Bundesen, C., Habekost, T., 2014. Theory of visual attention, in: Nobre, A., Kastner, S. (Eds.), *The Oxford handbook of attention*. Oxford University Press, pp. 1095–1121. doi:10.1093/oxfordhb/9780199675111.001.0001.
- Castles, A., 1996. Cognitive Correlates of Developmental Surface Dyslexia: A Single Case Study. *Cognitive Neuropsychology* 13, 25–50. doi:10.1080/026432996382051. publisher: Routledge _eprint: <https://doi.org/10.1080/026432996382051>.
- Castles, A., 2006. How does orthographic learning happen? From Inkmarks to Ideas: Current Issues in Lexical Processing , 151–179doi:10.4324/9780203841211-16.
- Castles, A., Rastle, K., Nation, K., 2018. Ending the Reading Wars: Reading Acquisition From Novice to Expert - Anne Castles, Kathleen Rastle, Kate Nation, 2018. *Psychological Science* 19, 5–51. doi:10.1177/1529100618772271.
- Chaffin, R., Morris, R.K., Seely, R.E., 2001. Learning new word meanings from context: A study of eye movements. *Journal of Experimental Psychology: Learning, Memory and Cognition* 27, 225–235. doi:10.1037//0278-7393.27.1.225.
- Chan, K.S.C., Yeung, P.S., 2020. Prediction of Chinese Reading Fluency by Verbal and Non-verbal Visual Attention Span Measures. *Frontiers in Psychology* 10. doi:10.3389/fpsyg.2019.03049. publisher: Frontiers.
- Chen, C., Schneps, M.H., Masyn, K.E., Thomson, J.M., 2016. The effects of visual attention span and phonological decoding in reading comprehension in dyslexia: A path analysis. *Dyslexia* 22, 322–344. doi:10.1002/dys.1543.
- Chetail, F., 2017. What do we do with what we learn? Statistical learning of orthographic regularities impacts written word processing. *Cognition* 163, 103–120. doi:10.1016/j.cognition.2017.02.015.
- Cunningham, A.E., 2006. Accounting for children's orthographic learning while reading text: Do children self-teach? *Journal of Experimental Child Psychology* 95, 56–77. doi:10.1016/j.jecp.2006.03.008.
- Cunningham, A.E., Perry, K.E., Stanovich, K.E., Share, D.L., 2002. Orthographic learning during reading: examining the role of self-teaching. *Journal of Experimental Child Psychology* 82, 185–199. doi:10.1016/S0022-0965(02)00008-5.
- De Luca, M., Borrelli, M., Judica, A., Spinelli, D., Zoccolotti, P., 2002. Reading Words and Pseudowords: An Eye Movement Study of Developmental Dyslexia. *Brain and Language* 80, 617–626. doi:10.1006/brln.2001.2637.
- Dubois, M., Kyllingsbaek, S., Prado, C., Musca, S.C., Peiffer, E., Lassus-Sangosse, D., Valdois, S., 2010. Fractionating the multi-character processing deficit in developmental dyslexia: Evidence from two case studies. *Cortex* 46, 717–738. doi:10.1016/j.cortex.2009.11.002. publisher: Elsevier.
- Engbert, R., Longtin, A., Kliegl, R., 2002. A dynamical model of saccade generation in reading based on spatially distributed lexical processing. *Vision research* 42, 621–636. doi:10.1016/S0042-6989(01)00301-7.
- Frey, A., Bosse, M.L., 2018. Perceptual span, visual span, and visual attention span: Three potential ways to quantify limits on visual processing during reading. *Visual Cognition* 26, 412–429. doi:10.1080/13506285.2018.1472163.
- Frith, U., 1985. Beneath the surface of developmental dyslexia.
- Germano, G.D., Reilhac, C., Capellini, S.A., Valdois, S., 2014. The phonological and visual basis of developmental dyslexia in Brazilian Portuguese reading children. *Frontiers in Psychology* 5, 1169. doi:10.3389/fpsyg.2014.01169.
- Ginestet, E., 2019. Modélisation bayésienne et étude expérimentale du rôle de l'attention visuelle dans l'acquisition des connaissances lexicales orthographiques. Ph.D. thesis. Université Grenoble Alpes.
- Ginestet, E., Phénix, T., Diard, J., Valdois, S., 2019. Modeling the length effect for words in lexical decision: The role of visual attention. *Vision Research* 159, 10–20. doi:10.1016/j.visres.2019.03.003.
- Ginestet, E., Valdois, S., Diard, J., 2022. Probabilistic modeling of orthographic learning based on visuo-attentional dynamics. *Psychonomic Bulletin & Review* .

- Ginestet, E., Valdois, S., Diard, J., Bosse, M.L., 2020. Orthographic learning of novel words in adults: effects of exposure and visual attention on eye movements. *Journal of Cognitive Psychology* 32, 785–804. doi:10.1080/20445911.2020.1823987. publisher: Routledge_eprint: <https://doi.org/10.1080/20445911.2020.1823987>.
- Grainger, J., Dufau, S., Montant, M., Ziegler, J.C., Fagot, J., 2012. Orthographic Processing in Baboons. *Science* 336, 245–248. doi:10.1126/science.1218152. publisher: American Association for the Advancement of Science Section: Report.
- Gurobi Optimization, LLC, 2021. Gurobi Optimizer Reference Manual.
- Habekost, T., 2015. Clinical tva-based studies: A general overview. *Frontiers in Psychology* 6, 290. doi:10.3389/fpsyg.2015.00290.
- Harris, T., Yang, Z., Hardin, J.W., 2012. Modeling underdispersed count data with generalized Poisson regression. *The Stata Journal* 12, 736–747. doi:10.1177/1536867X1201200412. publisher: SAGE Publications.
- Hautala, J., Hyönä, J., Aro, M., 2011. Dissociating spatial and letter-based word-length effects observed in readers' eye movement patterns. *Vision Research* 51, 1719–1727. doi:10.1016/j.visres.2011.05.015.
- Howard, D., 1996. Developmental Phonological Dyslexia: Real Word Reading Can Be Completely Normal. *Cognitive Neuropsychology* 13, 887–934. doi:10.1080/026432996381854.
- Huang, C., LOrusso, M.L., Luo, Z., Zhao, J., 2019. Developmental differences in the relationship between visual attention span and chinese reading fluency. *Frontiers in Psychology* 10, 2450. doi:10.3389/fpsyg.2019.02450.
- Joseph, H., Nation, K., 2018. Examining incidental word learning during reading in children: The role of context. *Journal of Experimental Child Psychology* 166, 190–211. doi:10.1016/j.jecp.2017.08.010.
- Joseph, H.S., Liversedge, S.P., Blythe, H.I., White, S.J., Rayner, K., 2009. Word length and landing position effects during reading in children and adults. *Vision Research* 49, 2078–2086. doi:10.1016/j.visres.2009.05.015.
- Joseph, H.S.S.L., Wonnacott, E., Forbes, P., Nation, K., 2014. Becoming a written word: Eye movements reveal order of acquisition effects following incidental exposure to new words during silent reading. *Cognition* 133, 238–248. doi:10.1016/j.cognition.2014.06.015.
- Juphard, A., Carbonnel, S., Valdois, S., 2004. Length effect in reading and lexical decision: Evidence from skilled readers and a developmental dyslexic participant. *Brain and Cognition* 55, 332–340. doi:10.1016/j.bandc.2004.02.035.
- Kliegl, R., Grabner, E., Rolfs, M., Engbert, R., 2004. Length, frequency, and predictability effects of words on eye movements in reading. *European journal of cognitive psychology* 16, 262–284. doi:10.1080/09541440340000213.
- Kwok, R.K., Ellis, A.W., 2014. Visual word learning in adults with dyslexia. *Frontiers in Human Neuroscience* 8, 264. doi:10.3389/fnhum.2014.00264.
- LaBerge, D., Samuels, S.J., 1974. Toward a theory of automatic information processing in reading. *Cognitive Psychology* 6, 293–323. doi:10.1016/0010-0285(74)90015-2.
- Lelonkiewicz, J.R., Ktori, M., Crepaldi, D., 2020. Morphemes as letter chunks: Discovering affixes through visual regularities. *Journal of Memory and language* 115, 104152.
- Loberg, O., Hautala, J., Hämäläinen, J.A., Leppänen, P.H., 2019. Influence of reading skill and word length on fixation-related brain activity in school-aged children during natural reading. *Vision Research* 165, 109–122. doi:10.1016/j.visres.2019.07.008.
- Lobier, M., Dubois, M., Valdois, S., 2013. The Role of Visual Processing Speed in Reading Speed Development. *PLOS ONE* 8, e58097. doi:10.1371/journal.pone.0058097. publisher: Public Library of Science.
- Lobier, M., Peyrin, C., Le Bas, J.F., Valdois, S., 2012. Pre-orthographic character string processing and parietal cortex: A role for visual attention in reading. *Neuropsychologia* 50, 2195–2204. doi:10.1016/j.neuropsychologia.2012.05.023.
- Lowell, R., Morris, R.K., 2014. Word length effects on novel words: Evidence from eye movements. *Attention, Perception and Psychophysics* 76, 179–189. doi:10.3758/s13414-013-0556-4.
- Maloney, E., Risko, E.F., O'Malley, S., Besner, D., 2014. Tracking the transition from sublexical to lexical processing: On the creation of orthographic and lexical representations. *The Quarterly Journal of Experimental Psychology* 62, 858–867. doi:10.1080/17470210802578385.
- Manis, F.R., 1985. Acquisition of word identification skills in normal and disabled readers. *Journal of Educational Psychology* 77, 78–90. doi:10.1037/0022-0663.77.1.78. place: US Publisher: American Psychological Association.
- Marinelli, C.V., Romani, C., Burani, C., McGowan, V.A., Zoccolotti, P., 2016. Costs and Benefits of Orthographic Inconsistency in Reading: Evidence from a Cross-Linguistic Comparison. *PLOS ONE* 11, e0157457. doi:10.1371/journal.pone.0157457. publisher: Public Library of Science.
- Marinelli, C.V., Zoccolotti, P., Romani, C., 2020. The ability to learn new written words is modulated by language orthographic consistency. *PLOS ONE* 15, e0228129. URL: <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0228129>, doi:10.1371/journal.pone.0228129. publisher: Public Library of Science.
- Martens, V.E., de Jong, P.J., 2008. Effects of repeated reading on the length effect in word and pseudoword reading. *Journal of Research in Reading* 31, 40–54. doi:10.1111/j.1467-9817.2007.00360.x.
- McClelland, J.L., Rumelhart, D.E., 1981. An interactive activation model of context effects in letter perception: I. An account of basic findings. *Psychological Review* 88, 375–407. doi:10.1037/0033-295X.88.5.375. place: US Publisher: American Psychological Association.
- McDonald, S.A., 2006. Effects of number-of-letters on eye movements during reading are independent from effects of spatial word length. *Visual Cognition* 13, 89–98. doi:10.1080/13506280500143367.
- Nation, K., Angell, P., Castles, A., 2007. Orthographic learning via self-teaching in children learning to read English: Effects of exposure, durability, and context. *Journal of Experimental Child Psychology* 96, 71–84. doi:10.1016/j.jecp.2006.06.004.
- Niolaki, G., Vousden, J., Terzopoulos, A., Taylor, L., Sephton, S., Masterson, J., 2020. Predictors of single word spelling in English-speaking children: A cross-sectional study. *Journal of Research in Reading* 43, 1–20. doi:10.1111/1467-9817.12330.
- Pagan, A., Nation, K., 2019. Learning words via reading: Contextual diversity, spacing, and retrieval effects in adults. *Cognitive Science* 43, e12705. doi:10.1111/cogs.12705.
- Pelli, D.G., Tillman, K.A., Freeman, J., SU, M., Berger, T.D., Majaj, N.J., 2007. Crowding and eccentricity determine reading rate. *Journal of Vision* 7, 1–36. doi:10.1167/7.2.20.
- Pellicer-Sanchez, A., 2016. Incidental L2 vocabulary acquisition from and while reading. *Studies in Second Language Acquisition* 38, 97–130. doi:10.1017/S0272263115000224.
- Perry, C., Zorzi, M., Ziegler, J.C., 2019. Understanding Dyslexia Through Personalized Large-Scale Computational Models. *Psychological Science* 30, 386–395. doi:10.1177/095679761882354010.1177/0956797618823540.

- Peyrin, C., Démonet, J.F., N'Guyen-Morel, M.A., Le Bas, J.F., Valdois, S., 2011. Superior parietal lobule dysfunction in a homogeneous group of dyslexic children with a visual attention span disorder. *Brain and Language* 118, 128–138. doi:10.1016/j.bandl.2010.06.005.
- Phenix, T., 2018. Modélisation bayésienne algorithmique de la reconnaissance visuelle de mots et de l'attention visuelle. Ph.D. thesis. Université Grenoble Alpes.
- Phenix, T., Valdois, S., Diard, J., 2018. Reconciling opposite neighborhood frequency effects in lexical decision: Evidence from a novel probabilistic model of visual word recognition, in: Rogers, T., Rau, M., Zhu, X., Kalish, C.W. (Eds.), *Proceedings of the 40th Annual Conference of the Cognitive Science Society*, Cognitive Science Society, Austin, TX. pp. 2238–2243.
- Prado, C., Dubois, M., Valdois, S., 2007. The eye movements of dyslexic children during reading and visual search: Impact of the visual attention span. *Vision Research* 47, 2521–2530. doi:10.1016/j.visres.2007.06.001.
- Pritchard, S.C., Coltheart, M., Marinus, E., Castles, A., 2018. A Computational Model of the Self-Teaching Hypothesis Based on the Dual-Route Cascaded Model of Reading. *Cognitive Science* 42, 722–770. doi:10.1111/cogs.12571.
- Provazza, S., Adams, A.M., Giolfè, D., Roberts, D.J., 2019. Double trouble: visual and phonological impairments in English dyslexic readers. *Frontiers in psychology* 10, 2725. doi:10.3389/fpsyg.2019.02725. publisher: Frontiers.
- R Core Team, 2020. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.
- Rayner, K., 1998. Eye Movements in Reading and Information Processing: 20 Years of Research. *Psychological Bulletin* 124, 372–422. doi:10.1037/0033-2909.124.3.372.
- Rayner, K., Sereno, S.C., Raney, G.E., 1996. Eye movement control in reading: A comparison of two types of models. *Journal of Experimental Psychology: Human perception and performance* 22, 1188–1200. doi:10.1037/0096-1523.22.5.1188.
- Reichle, E.D., Rayner, K., Pollatsek, A., 2003. The E-Z Reader model of eye-movement control in reading: Comparisons to other models. *Behavioral and brain sciences* 26, 445–476. doi:10.1017/S0140525X03000104.
- Reilhac, C., Peyrin, C., Démonet, J.F., Valdois, S., 2013. Role of the superior parietal lobules in letter-identity processing within strings: Fmri evidence from skilled and dyslexic readers. *Neuropsychologia* 51, 601–612. doi:10.1016/j.neuropsychologia.2012.12.010.
- Saghiran, A., Valdois, S., Diard, J., 2020. Simulating length and frequency effects across multiple tasks with the Bayesian model BRAID-Phon, in: 42nd Annual Virtual Meeting of the Cognitive Science Society, Toronto, Canada. pp. 3158–3163.
- Scarf, D., Boy, K., Reinert, Anelise Uberand Devine, J., Güntürkün, O., Colombo, M., 2016. Orthographic processing in pigeons (*columba livia*). *Proceedings of the National Academy of Science* 113, 11272–11276. doi:10.1111/1467-9817.12043.
- Share, D.L., 1995. Phonological recoding and self-teaching: sine qua non of reading acquisition. *Cognition* 55, 151–218. doi:10.1016/0010-0277(94)00645-2.
- Share, D.L., 1999. Phonological Recoding and Orthographic Learning: A Direct Test of the Self-Teaching Hypothesis. *Journal of Experimental Child Psychology* 72, 95–129. doi:10.1006/jecp.1998.2481.
- Share, D.L., Shalev, C., 2004. Self-teaching in normal and disabled readers. *Reading and Writing* 17, 769–800. doi:10.1007/s11145-004-2658-9.
- Spinelli, D., de Luca, M., di Filippo, G., Mancini, M., Martelli, M., Zoccolotti, P., 2005. Length effect in word naming in reading: Role of reading experience and reading deficit in Italian readers. *Developmental Neuropsychology* 27, 217–235. doi:10.1207/s15326942dn2702_2.
- Stefanac, N., Spencer-Smith, M., Méadhbh, B., Vangkilde, S., Castles, A., Bellgrove, M., 2019. Visual processing speed as a marker of immaturity in lexical but not sublexical dyslexia. *Cortex* 120, 567–581. doi:10.1016/j.cortex.2019.08.004.
- Stenneken, P., Egetemeir, J., Schulte-Körne, G., Müller, H.J., Schneider, W.X., Finke, K., 2011. Slow perceptual processing at the core of developmental dyslexia: A parameter-based assessment of visual attention. *Neuropsychologia* 49, 3454–3465. doi:10.1016/j.neuropsychologia.2011.08.021.
- Suarez-coalla, P., Alvarez-Cañizo, M., Cuetos, F., 2016. Orthographic learning in Spanish children. *Journal of Research in Reading* 39, 292–311. doi:10.1111/1467-9817.12043.
- Suárez-Coalla, P., Ramos, S., Álvarez Cañizo, M., Cuetos, F., 2014. Orthographic learning in dyslexic Spanish children. *Annals of Dyslexia* 64, 166–181. doi:10.1007/s11881-014-0092-5.
- Townsend, J.T., 1971. Alphabetic confusion: A test of models for individuals. *Perception & Psychophysics* 9, 449–454. doi:10.3758/BF03208950.
- Tucker, R., Castles, A., Laroche, A., Deacon, S.H., 2016. The nature of orthographic learning in self-teaching: Testing the extent of transfer. *Journal of Experimental Child Psychology* 145, 79–94. doi:10.1016/j.jecp.2015.12.007.
- Valdois, S., 2022. The visual attention span deficit in developmental dyslexia: Review of evidence for a visual-attention-based deficit. *Dyslexia* doi:10.1002/dys.1724.
- Valdois, S., Bidet-Ildéi, C., Lassus-Sangosse, D., Reilhac, C., N'guyen-Morel, M.A., Guinet, E., Orliaguet, J.P., 2011. A visual processing but no phonological disorder in a child with mixed dyslexia. *Cortex* 47, 1197–1218. doi:10.1016/j.cortex.2011.05.011.
- Valdois, S., Bosse, M.L., Ans, B., Carbonnel, S., Zorman, M., David, D., Pellat, J., 2003. Phonological and visual processing deficits can dissociate in developmental dyslexia: Evidence from two case studies. *Reading and Writing* 16, 541–572. doi:10.1023/A:1025501406971.
- Valdois, S., Bosse, M.L., Tainturier, M.J., 2004. The cognitive deficits responsible for developmental dyslexia: Review of evidence for a selective visual attentional disorder. *Dyslexia* 10, 339–363. doi:10.1002/dys.284.
- Valdois, S., Phenix, T., Fort, M., Diard, J., 2021a. Atypical viewing position effect in developmental dyslexia: A behavioural and modelling investigation. *Cognitive Neuropsychology* doi:10.1080/02643294.2021.2004107.
- Valdois, S., Reilhac, C., Ginestet, E., Bosse, M.L., 2021b. Varieties of cognitive profiles in poor readers: Evidence for a vas-impaired subtype. *Journal of Learning Disabilities* 54, 221–233. doi:10.1177/0022219420961332.
- Valdois, S., Roulin, J.L., Bosse, M.L., 2019. Visual attention modulates reading acquisition. *Vision Research* 165, 152–161. doi:10.1016/j.visres.2019.10.011.
- van Viersen, S., Protopapas, A., Georgiou, G.K., Parila, R., Ziaka, L., de Jong, P., 2022. Lexicality effects on orthographic learning in beginning readers and advanced readers of Dutch: An eye-tracking study. *Quarterly Journal of Experimental Psychology* 75, 1135–1154. doi:10.1177/17470218211047420.
- Vitu, F., O'Regan, J., Mittau, M., 1990. Optimal landing position in reading isolated words and continuous text. *Perception & Psychophysics* 47, 583–600. doi:10.3758/BF03203111.
- Waechter, S., Besner, D., Stolz, J.A., 2011. Basic processes in reading: Spatial attention as a necessary preliminary to orthographic and semantic processing. *Visual Cognition* 19, 171–202. doi:10.1080/13506285.2010.517228.

- Williams, R., Morris, R., 2004. Eye movements, word familiarity and vocabulary acquisition. *European Journal of Cognitive Psychology* 16, 312–339. doi:10.1080/09541440340000196.
- Zhao, J., Liu, M., Miu, H., Huang, C., 2018. The visual attention span deficit in chinese children with reading fluency difficulty. *Research in Developmental Disabilities* 73, 76–86. doi:10.1016/j.ridd.2017.12.017.
- Ziegler, J.C., Perry, C., Zorzi, M., 2014. Modelling reading development through phonological decoding and self-teaching: implications for dyslexia. *Phil. Trans. R. Soc. Biological Sciences* 369, 20120397. doi:10.1098/rstb.2012.0397.
- Zoccolotti, P., De Luca, M., Di Pace, E., Gasperini, F., Judica, A., Spinelli, D., 2005. Word length effect in early reading and in developmental dyslexia. *Brain and Language* 93, 369–373. doi:10.1016/j.bandl.2004.10.010.
- Zoubinetzky, R., Bielle, F., Valdois, S., 2014. New insights on developmental dyslexia subtypes: Heterogeneity of mixed reading profiles. *PloS ONE* 9, e99337. doi:10.1371/journal.pone.0099337.

Appendix A. Stimuli

4-letter words: acme, arak, ares, barb, bess, boon, brig, cell, chin, coup, dade, deer, dill, dyne, enos, gale, gaud, gent, hemp, joss, june, kivu, lear, leek, loch, buri, cony, lura, mali, marr, mink, moth, nara, huns, oath, peru, quod, role, rook, scut, slat, soul, tarn, tofu, topi, tosh, tree, vial, womb, yeas, aide, ainu, aryl, attu, oleg, bert, body, buna, byes, caff, capn, miry, dodd, dram, edam, feat, feds, fogg, ludo, fore, gogo, gown, grot, grub, hake, hume, husk, koan, lakh, pron, menu, mort, nett, orly, oxen, pane, pomp, quay, sham, sims, skit, talc, togs, tory, vail, vats, volt, weft, wold, yule

5-letter words: arabs, aroma, aspen, babel, baker, balsa, berry, blues, cache, chump, codex, compo, crust, dicks, dildo, flank, drake, fanny, dolly, greer, harem, horne, jonah, keane, lewis, loren, macon, males, maple, oasis, ozone, pansy, penis, photo, rabbi, clasp, rotor, rover, rumba, skull, sloan, snack, syrup, tamil, teeth, toque, trier, uncle, vigil, wayne, anvil, aorta, argos, asp, atoll, attic, aught, blood, bourn, canoe, carey, chris, cleva, della, dinar, ernie, ether, folio, foyer, gibby, gusto, heron, highs, ivory, jones, katie, kurus, levin, maine, navvy, rhode, robot, sabra, sadie, saran, scuba, sewer, shank, sioux, skiff, slush, spoof, sprig, swath, toska, twine, walls, weiss, whorl, wilde

6-letter words: ablaut, anklet, arrack, beeves, borage, centum, cicala, cotman, cowmen, czechs, dalton, dowser, flagon, gigolo, hotpot, howdah, icemen, kronor, krutch, kummel, lugger, mender, noshes, office, oxcart, pignut, poppet, ranker, rioter, sacker, sateen, scrota, seekin, shensi, stamen, street, sundew, tatian, tibiae, tomtom, torrio, tumuli, xavier, yeoman, yogurt, yonder, zenith, zephyr, zinnia, zombie, andrus, beirut, bistro, bustle, cactus, cartel, catgut, chukka, cicero, delvin, dibble, doddle, duenna, dustup, emblem, escudo, family,

friend, fulmar, gasmen, gooier, guizot, hangup, hannah, hippie, hopper, howell, idiocy, jasper, lemons, newton, orgasm, persia, pulsar, quincy, rapist, rogues, rotter, runnel, sayers, schulz, sidney, sinker, strang, strata, varian, volume, wicket, wilson, yokuts

7-letter words: affaire, alumnae, anthill, autarky, barnaba, blanket, blemish, brooder, buildup, clayton, colonus, waiving, corrals, country, crystal, dawdler, decoder, divider, doublet, dresser, economy, egerton, erosion, evasion, firearm, flyways, francis, gingham, gouache, goulash, grenada, hormone, imagery, inkling, longbow, macedon, maurice, nemesis, newport, newsmen, oregano, panoply, pedicel, poussin, prowess, referee, seaport, stratum, virgule, vulture, antenna, babcock, beaches, bloomer, booklet, buttock, cabbage, calypso, concept, dilemma, diploma, dorothy, forrest, garrett, gazelle, gestapo, grafton, heckler, heywood, jackson, jenkins, lincoln, liqueur, luggage, mailman, mankind, mongrel, neilson, oranges, pattern, phantom, pitcher, pitfall, pointer, pompano, pretext, privacy, provost, sangria, schmidt, siberia, slipper, snowman, stinger, surgery, syrians, tremolo, untruth, valerie, virgins,

8-letter words: besieger, bombsite, bootlace, bullhide, cajolery, causerie, clifford, decoking, division, entresol, eyetooth, families, findsome, fireclay, gallants, glumness, gripsack, icefloes, infamies, lifebelt, lifebuoy, lummoxes, majority, mastoids, medicine, orchises, overplus, parterre, prattler, property, psalmody, putsches, quirinal, raciness, raillery, rankness, rockhall, tenpence, throstle, tidemark, toadyism, tollgate, transfer, turnspit, wigmaker, wineskin, wiriness, yugoslav, zeppelin, zimbabwe, addendum, botulism, boutique, bulgaria, cambodia, cassette, causeway, churches, commando, compiler, cupboard, deathbed, detritus, eyepiece, finisher, haitians, handbook, heraldry, holiness, ideology, instance, laxative, licensee, machismo, metaphor, musician, namesake, nebraska, plastics, pretense, proposal, roadster, rushmore, seedling, sherlock, softness, specimen, speeches, stimulus, tamarind, tasmmania, tendency, theology, treasury, ugliness, universe, werewolf, westwood, winfield, woodside

9-letter words: ablatives, australia, blowflies, blutwurst, bourguiba, bowerbird, bridewell, cominform, companies, contriver, costumier, crimplene, cuckoldry, deauville, exhusband, flageolet, flashcube, abasement, fortifier, identikit, lobscouse, lowlander, lowliness, luckiness, lumbermen, luridness, lustiness, mistiness, moralizer, newspaper, nunneries, oratories, orrisroot, patricide, phagocyte, phalanges, polyether, punctilio, repletion, sandshoes, scenarist, september, sixtieths, smoochers, stridence, sunniness, technique, timid-

ness, treatment, woodlouse, agreement, attention, candidate, cerebella, charabanc, charwoman, chiseller, cicatrice, developer, diathesis, driveller, duchesses, fooleries, forcemeat, forewoman, garrulity, germicide, gushiness, hothouses, ignorance, lactation, lazaruses, leucotomy, materials, noctiluca, obscurant, omnibuses, orangeade, packhorse, panatella, papyruses, peccaries, penknives, personnel, plasterer, poltroons, stokehold, striation, sucklings, suffusion, sulkiness, sunfishes, tailboard, telltales, territory, tigresses, wesleyans, youngster, zimmerman, zoologist

10-letter words: andromache, basketball, burckhardt, burlesques, categories, coagulants, conception, concretion, conversion, coronaries, corrigenda, crustiness, delphinium, employment, evaluation, flagellant, gingersnap, graphology, hobbyhorse, horseflesh, intactness, keypunches, lordliness, maidenhood, manageress, mortuaries, newsletter, pliability, postscript, preclusion, preference, properties, propionate, psychology, quintuplet, saleswomen, savageness, scrollwork, specialist, speleology, stonemason, submission, suspension, telephotos, terramycin, thrashings, threepence, truculence, undulation, vulgarians, alpenstock, anglomania, anointment, antiheroes, apoplexies, artfulness, assumption, bestiaries, braininess, businesses, clerestory, collieries, colloquies, conclusion, conference, dishabille, eisteddfod, foundation, giantesses, glossiness, goldfishes, hibiscuses, homoeopath, horselaugh, horsewoman, husbandman, industries, instrument, intendants, inwardness, irishwoman, mainstream, minuteness, parliament, petrolatum, preferment, presbytery, psalteries, reputation, resolution, rheumatics, scantlings, subsidizer, succulence, supplanter, swordstick, throughway, waterpower, workpeople, yellowness