



**HAL**  
open science

## The influence of word frequency on word reading speed when individuals with macular diseases read text

Natacha Stolowy, Aurelie Calabrese, Lauren Sauvan, Carlos Aguilar, Thomas François, Núria Gala, Frédéric Matonti, Eric Castet

► **To cite this version:**

Natacha Stolowy, Aurelie Calabrese, Lauren Sauvan, Carlos Aguilar, Thomas François, et al.. The influence of word frequency on word reading speed when individuals with macular diseases read text. Vision Research, 2019, 155, pp.1-10. 10.1016/j.visres.2018.12.002 . hal-02360849

**HAL Id: hal-02360849**

**<https://hal.science/hal-02360849v1>**

Submitted on 19 Oct 2021

**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.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

**Title: The influence of word frequency on word reading speed when individuals with macular diseases read text.**

**Short title:** Word-frequency effects in macular diseases

**Authors:** Natacha Stolowy<sup>a,\*</sup>, Aurélie Calabrèse<sup>b,c,\*</sup>, Lauren Sauvan<sup>a</sup>, Carlos Aguilar<sup>d,e</sup>, Thomas François<sup>f,g</sup>, Núria Gala<sup>b,h</sup>, Frédéric Matonti<sup>a,b,i</sup>, Eric Castet<sup>b,c</sup>

**Affiliations:**

<sup>a</sup> North Hospital, Marseille France

<sup>b</sup> Aix-Marseille University, Marseille, France

<sup>c</sup> Laboratoire de Psychologie Cognitive, CNRS, Marseille, France

<sup>d</sup> Nice Sophia-Antipolis University, Nice, France

<sup>e</sup> Laboratoire Bases Corpus Langage, CNRS, Nice, France

<sup>f</sup> Université catholique de Louvain

<sup>g</sup> Institut Langage et Communication, FNRS, Louvain, Belgique

<sup>h</sup> Laboratoire Parole et Langage, Aix-en-Provence, France

<sup>i</sup> University Hospital of La Timone, Marseille, France

\* N.S. and A.C. contributed equally to this work.

**Corresponding author:**

Aurélie Calabrèse (AC)

aurelie.calabrese@univ-amu.fr

3 place Victor Hugo - 13003 Marseille, France

**Word count:**

Abstract = 213 words

Text = 5092 words

*Intro = 1095; Methods = 1322; Results = 1001; Discussion = 1535; Conclusion = 197*

References = 70; Figures = 4; Tables = 3

**Funding:**

This work was supported by the Fondation de France and by the Belgian FNRS.

**Declarations of interest:**

None

1 **ABSTRACT**

2 People with central field loss (CFL) use peripheral vision to identify words. Eccentric vision  
3 provides ambiguous visual inputs to the processes leading to lexical access. Our purpose was  
4 to explore the hypothesis that this ambiguity leads to strong influences of inferential  
5 processes, our prediction being that increasing word frequency would decrease word reading  
6 time. Individuals with bilateral CFL induced by macular diseases read French sentences  
7 displayed with a self-paced reading method. Reading time of the last word of each sentence  
8 (target word) was recorded. Each target word (in sentence n) was matched with a synonym  
9 word (in sentence n+1) of the same length. When using absolute frequency value (Analysis  
10 1), we found that reading time of target words decreased when word frequency increases,  
11 even when controlling for word length. The amplitude of this effect is larger than reported in  
12 previous investigations of reading with normal subjects. When comparing the effect of  
13 relative frequency (low vs. high) within each pair of synonyms (Analysis 2), results show the  
14 same pattern as the one observed in Analysis 1. Our results demonstrate clear-cut frequency  
15 effects on word reading time and suggest that inferential processes are stronger in CFL  
16 readers than in normally sighted observers. These results might also help design text  
17 simplification tools tailored for low-vision patients.

18

19

20 **KEYWORDS**

21 central field loss, macular degeneration, reading, psycholinguistics, word frequency, lexical  
22 simplification

23

24

25

26           **1. INTRODUCTION**

27

28   Age-related Macular Degeneration (AMD) is one of the main causes of low vision in  
29   industrialized countries and affects about 20% of the population over 65 years (Klein et al.,  
30   1999; Wong et al., 2014). Like other macular diseases such as Stargardt’s disease, cone  
31   dystrophy or myopic maculopathy, AMD causes a central field loss (CFL) that severely  
32   impairs reading, an activity that is usually the main complaint of these patients (Brown et al.,  
33   2014; Elliott et al., 1997; Mangione et al., 1998; Wolffsohn & Cochrane, 1999).

34

35   Currently-available treatments for AMD cannot always restore central vision and, in that case,  
36   patients’ support is mainly based on low-vision rehabilitation and the use of visual aids  
37   (Meyniel et al., 2017). However, the use of these visual aids, whose principles are essentially  
38   based on magnification, is not sufficient to restore reading performance in patients with AMD  
39   (Chung et al., 1998; Legge, 2007; Li et al., 2002). This is why many efforts have been made  
40   to improve reading speed in peripheral vision by manipulating other visual factors such as: (1)  
41   text presentation mode, e.g. static, scrolling or flashing one word at a time (Bowers et al.,  
42   2004; Fine & Peli, 1998; Rubin & Turano, 1994); (2) characteristics of the text, e.g. font  
43   (Bernard et al., 2016; Bragg et al., 2017; Mansfield et al., 1996), text orientation (Calabrèse et  
44   al., 2017), letter, word or interline spacing (Bernard et al., 2007; Blackmore-Wright et al.,  
45   2013; Calabrèse et al., 2010; Chung, 2002; Chung et al., 2008; Yu et al., 2007); (3) perceptual  
46   learning, i.e. the performance improvement that follows practice of a sensory task  
47   (Chung et al., 2004; **Chung, 2011**). Mixed results have been obtained in terms of the clinical  
48   efficiency of these manipulations.

49

50 In this paper, we explore another consequence for CFL patients of the necessity of using their  
51 eccentric vision. Words processed in the periphery provide a very ambiguous low-level visual  
52 input, due for instance to confusions between letters (Pelli, 2008), which should induce a high  
53 influence of lexical inference processes (Legge et al., 1997; Legge et al., 1997). Basically,  
54 when faced with an ambiguous percept (say a three-letter string starting with “c” and ending  
55 with “r”), a patient would make a guess about the most likely word having induced this  
56 percept (probably “car”). Thus, word frequency is expected to influence the speed with which  
57 a guess would be made. This is consistent with a recent investigation of eye movements in  
58 AMD patients during short-sentences reading that reported clusters of fixations that  
59 significantly reduce reading speed (Calabrèse et al., 2016). The clusters in some parts of a  
60 sentence indicated that specific words required dozens of fixations before successful  
61 identification, while the other ones were read with fewer fixations. The presence/absence of  
62 clusters was a significant predictor of reading speed, independently of the perceptual span size  
63 (Calabrèse et al., 2014). Interestingly, these clusters were more likely to occur in sentences  
64 containing a low-frequency word. This new finding suggested that low-frequency words  
65 might cause a local slowing of reading speed (i.e. a dense cluster of fixations) due to more  
66 difficult lexical access. However, the authors were not able to quantify the strength of the  
67 word frequency effect because their protocol did not allow for a clear determination of which  
68 word was being processed when a cluster occurred.

69

70 The main goal of the present work is to investigate this question further and study the direct  
71 effect of word frequency on word reading speed in individuals with bilateral macular disease.  
72 For normally sighted readers, factors such as word familiarity or word frequency are known to  
73 affect reading performance at different perceptual levels. To the best of our knowledge, the  
74 quantitative relationship between word frequency and natural reading performance with CFL

75 has not been investigated yet. This contrasts with the wealth of the literature for normally  
76 sighted observers (Clifton Jr. et al., 2016; Kliegl et al., 2006; Rayner, 1998). The main goal of  
77 the present study is thus to test whether word frequency affects word reading speed in  
78 individuals with CFL: it is expected that high-frequency words would lead to shorter reading  
79 times than low-frequency words. Two distinct analyses were performed to test the effect of  
80 word frequency on reading performance: Analysis 1, using absolute values of frequency (in  
81 occurrences/million) and Analysis 2, using frequency categories (low vs. high).

82

83 We used a reading presentation mode, known as “self-paced reading”, that has never been  
84 used before in low-vision reading studies. Self-paced reading has been extensively used in  
85 psycholinguistic studies as a proxy for natural reading (Aaronson & Scarborough, 1976; Just  
86 et al., 1982; Mitchell & Green, 1978). In this context, natural reading, also called eye-  
87 mediated reading, refers to the natural process of reading that is used in daily life and requires  
88 eye movements; this has to be contrasted with paradigms such as RSVP (Rapid Serial Visual  
89 Presentation) requiring few or no eye movements since all words are presented one at a time  
90 at the same position (Rubin & Turano, 1992). With the self-paced reading paradigm,  
91 sentences appear as a whole, with all words masked and revealed only one at a time.

92 Compared to other sequential word presentation modes, the self-paced reading paradigm  
93 reduces the amount of meaningful text to a single word, while maintaining the overall spatial  
94 layout of the sentence. Given that only one word is available to the reader at any given time,  
95 this paradigm makes it possible to identify in real time which word is being read, and thus to  
96 accurately measure the actual time spent reading each word.

97

98 Our secondary goal was to introduce the technique of text simplification and its potential  
99 benefit to facilitate reading performance with CFL. Text simplification is the process of

100 reducing the complexity of a text by modifying the vocabulary or the syntax, while preserving  
101 its informational content (Shardlow, 2014). The purpose of this technique is to transform a  
102 text into a different one, which will convey the same message, while being easier to read and  
103 understand by a broader audience (Saggion, 2017). Text simplification techniques have been  
104 applied to dyslexic and aphasic patients and have been shown to improve reading  
105 performance (Siddharthan, 2014). However, text simplification has not yet been considered  
106 for CFL patients. The present work constitutes a first step towards the investigation of the  
107 simplification significance for this specific population. By testing the effect of word  
108 frequency on individuals with CFL, we focused on a small part of text simplification: the  
109 lexical substitution. Our results on word frequency suggest that lexical simplification  
110 techniques in general could be used successfully as a new reading aid for individuals with  
111 CFL.

112

## 113 **2. METHODS**

114

### 115 **2.1 Participants**

116 29 participants were recruited from December 2016 to June 2017 from the Low-Vision Clinic  
117 of the La Timone Hospital (Marseille, France). Recruited patients had to present a bilateral  
118 central scotoma, accompanied or not by an additional peripheral lesion and be fluent French  
119 readers. Patients with ophthalmologic disorders other than maculopathy (e.g. glaucoma) or  
120 cognitive disorders were not included. There was no minimal visual acuity threshold. Among  
121 the 29 patients recruited, 6 were excluded from the study because they could not read an  
122 entire block of text due to their low-functional vision. 23 participants (14 females) were  
123 included in the present work. They had different etiologies: atrophic AMD (n = 11), exudative  
124 AMD (n = 2), Stargardt's disease (n = 4), cone dystrophy (n = 1) and myopic retinopathy (n =



125 5) (Kumar et al., 2017; Silva, 2012; Vongphanit et al., 2002). Mean age ( $\pm$  SD) was  $72 \pm 15$   
126 years and ranged from 32 to 88. Mean best-corrected visual acuity was **0.87 logMAR**.  
127 Additional information was collected for each individual: lens status, duration of  
128 maculopathy, low-vision rehabilitation history and reading habits. The research was approved  
129 by the Ethics Committee of the French Society of Ophthalmology (IRB 00008855 Société  
130 Française d’Ophtalmologie IRB#1) and carried out in accordance with the Code of Ethics of  
131 the World Medical Association (Declaration of Helsinki). Informed consent was obtained  
132 from all participants after complete explanation of the nature and possible consequences of  
133 the study. Table 1 summarizes the participants’ demographic, visual and reading  
134 characteristics.  
135

Partici pants	Gen der	Age (years)	Etiology	Disea se onset (years)	Low-vision rehabi litatio n	Visual acuity (logMAR)	Lens status	Daily reading time (min/day)
1	M	69	Atrophic AMD	9	Yes	0.5	Phakic	0
2	M	76	Atrophic AMD	2	No	1	Phakic	0
3	F	76	Atrophic AMD	16	Yes	1	Phakic	0
4	F	77	Atrophic AMD	7	No	NA	Pseudophakic	45
5	F	79	Atrophic AMD	5	Yes	0.8	Phakic	5
6	F	80	Atrophic AMD	6	Yes	1	Phakic	25
7	M	81	Atrophic AMD	2	Yes	0.8	Pseudophakic	15
8	F	82	Atrophic AMD	12	Yes	0.5	Phakic	10
9	F	83	Atrophic AMD	1	Yes	0.7	Pseudophakic	0
10	F	85	Atrophic AMD	15	No	0.7	Pseudophakic	60
11	M	88	Atrophic AMD	22	Yes	0.6	Phakic	NA
12	M	85	Exudative AMD	5	Yes	0.7	Pseudophakic	60
13	F	87	Exudative AMD	1.5	Yes	0.5	Pseudophakic	20
14	M	47	Stargardt’s disease	24	No	1.2	Phakic	360
15	F	48	Stargardt’s disease	35	Yes	1.3	Phakic	20
16	M	51	Stargardt’s disease	37	No	1.2	Phakic	60
17	F	72	Stargardt’s disease	16	No	1	Phakic	120
18	F	32	Cone dystrophy	30	No	1	Phakic	NA
19	M	59	Myopic retinopathy	15	Yes	1	Phakic	3

20	F	69	Myopic retinopathy	21	Yes	1.3	Pseudophakic	120
21	F	70	Myopic retinopathy	62	Yes	0.8	Pseudophakic	12
22	M	78	Myopic retinopathy + AMD	6	Yes	0.6	Pseudophakic	5
23	F	82	Myopic retinopathy + AMD	41	Yes	1	Pseudophakic	30

136

137 **Table 1: Participants' demographic, visual and reading characteristics.**

138 **Visual acuity and lens status are given for the tested eye only. Participants with central scotoma**  
139 **only are highlighted in yellow. Participants with central scotoma and peripheral lesions are**  
140 **highlighted in blue. Participants with a reading time different from 0 were considered as “still**  
141 **reading” (cf Analyses 1 and 2). Missing values are represented by the symbol NA.**

142

## 143 **2.2 Apparatus and stimuli**

144 Stimuli were displayed on an LCD HP LE2201W monitor with a 60Hz screen refresh rate.  
145 Within the full display area of this monitor (47.4 x 29.6 cm; 1680 x 1050 pixels), a 1400 x  
146 1050 pixel window was used to display our stimuli. At the viewing distance of 40 cm, this  
147 window subtended 56° x 42°. Stimuli were created with the PsychoPy library (Peirce, 2007;  
148 Peirce, 2009). Characters were black on a white background. Sentences were aligned to the  
149 left and displayed in Courier (non proportional font). Print size was chosen optimally for each  
150 participant as the value of his/her Critical Print Size, measured before testing with a French  
151 computerized version of MNREAD (Calabrèse et al., 2014; Calabrèse et al., 2017). Reading  
152 was monocular (eye with better visual acuity) with an appropriate correction for near vision  
153 (wide-field Metrovision lenses).

154

## 155 **2.3 Reading material**

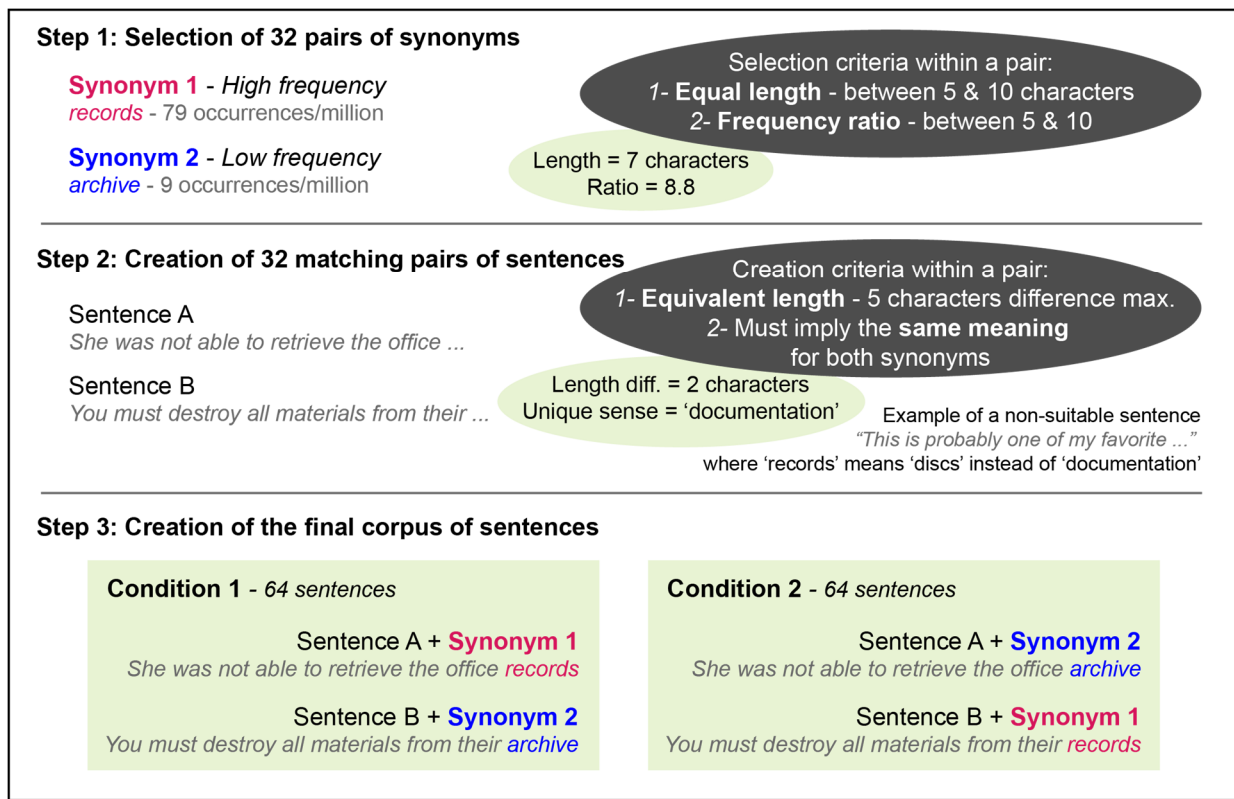
156 Reading material was created in French using ReSyf, a French graded lexicon (Billami et al.,  
157 2018) and Lexique3, a French lexical database (New et al., 2007). First, we created hundreds  
158 of pairs of synonyms by selecting a word of any frequency along with its higher frequency

159 synonym. We used the frequency of the lemma, which includes all the inflexional variants of  
160 a word (e.g. gender, number, tense, etc.). Lemma frequency (in occurrences / million) was  
161 estimated from a corpus of French film subtitles. Among the synonym pairs created, we then  
162 selected the ones that met the following criteria: (1) same number of characters within a pair  
163 (ranging from 5 to 10 characters); (2) frequency ratio between a high-frequency word and its  
164 low-frequency synonym comprised between 5 and 10. A total of 32 pairs of synonyms were  
165 selected (Figure 1 - Step 1).

166

167 From these 32 pairs of target words, we then created 32 pairs of matching sentences so that  
168 each word from a pair could fit at the end of either sentence of the corresponding sentence  
169 pair. To this end, pairs of sentences were specifically designed to fit the single common sense  
170 of a synonym pair. An example is given in Figure 1, with the polysemic word ‘records’, that  
171 means both ‘documentation’ and ‘discs’, and its synonym ‘archive’. In the sentence “This is  
172 probably one of my favorite records”, records and archive convey different senses and cannot  
173 be considered interchangeable synonyms anymore. Therefore, acceptable sentences for this  
174 synonym pair must refer to the meaning ‘documentation’ only. Apart from this context  
175 constraint, all sentences were created similar in length (41 to 65 characters), with a maximum  
176 difference of 5 characters within a pair (Figure 1 - Step 2). Lastly, we created our final  
177 reading material by combining these sentence pairs with their matching pairs of synonyms  
178 (Figure 1 - Step 3). In Condition 1, the first word of a pair was assigned to the first sentence  
179 of the corresponding pair, while the second word was assigned to the second sentence, thus  
180 creating 64 full sentences. In Condition 2, the “sentence – word” matching was reversed to  
181 create a different set of 64 full sentences. These two experimental conditions allowed us to  
182 counterbalance any potential effect of the sentence itself (structure, complexity, predictability)

183 by randomly assigning participants to Condition 1 or 2 (Steen-Baker et al., 2017). Figure 1  
 184 illustrates the full creation process.



185  
 186 **Figure 1: Creation of the reading material**

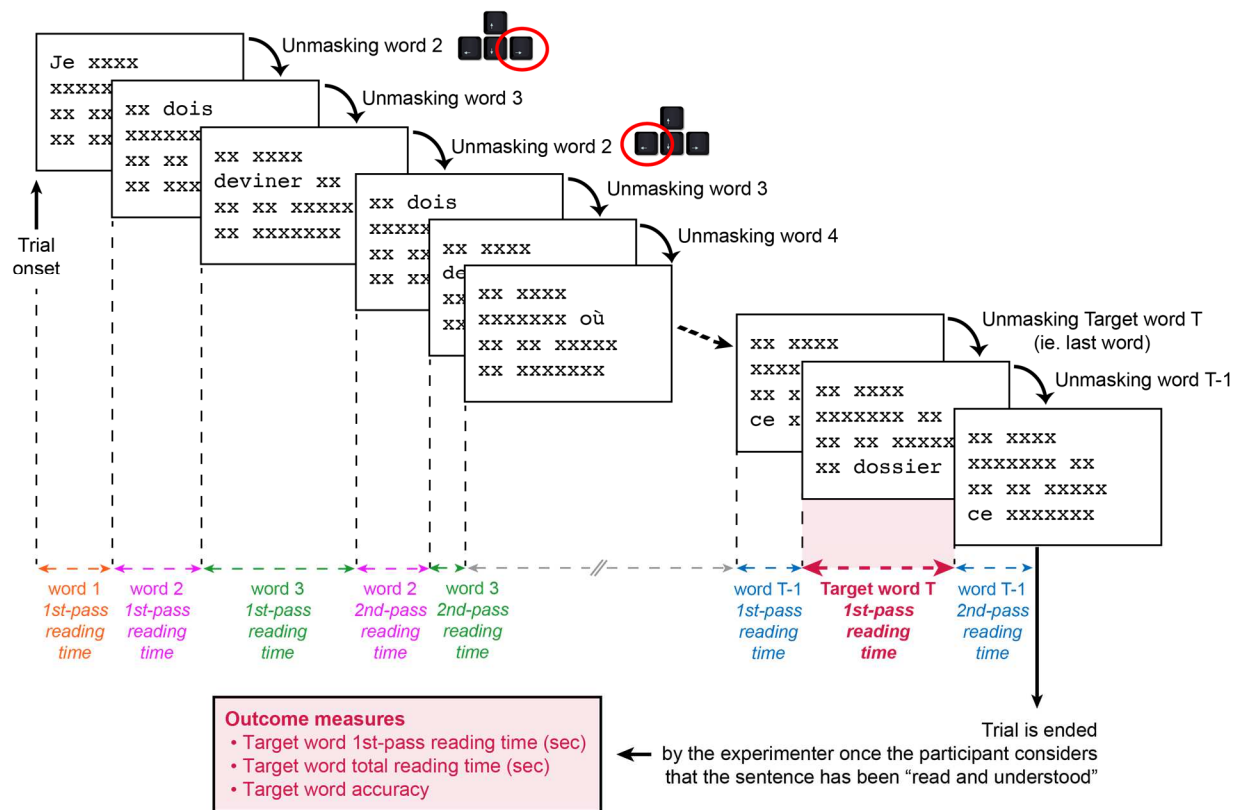
187 **Illustration of the reading material creation using the synonym pair “records / archive”. Both**  
 188 **words in this example contain 7 characters and have a frequency ratio (in French) of  $79 / 9 = 8.8$ .**  
 189 **Sentences A and B were created so that either word of this synonym pair can fit at the end of**  
 190 **both sentences. In Condition 1, sentence A contains the target word 1 (“records”) and sentence B**  
 191 **contains the target word 2 (“archive”). In Condition 2, sentence A contains the target word 2 and**  
 192 **sentence B contains the target word 1. Note that this example is given in English to ease**  
 193 **understanding. See Appendix 1 for the complete set of French sentences created.**

194  
 195 **2.4 Reading procedure and experimental design**

196 Sentences were presented within 4 blocks of 16 trials (8 pairs of sentences) each. Participants  
 197 were randomly assigned to Condition 1 or 2 and read between two and four blocks, depending  
 198 on their individual reading speed and level of fatigue. Sentences were displayed with the self-

199 paced reading paradigm and presented randomly within each block (Figure 2). Participants  
200 were instructed to read aloud as quickly and accurately as possible as soon as the sentence  
201 appeared. At the beginning of each trial, the experimenter triggered the display of a sentence,  
202 which appeared on the screen with the first word unmasked and all other words masked by  
203 strings of “x”. Participants were in charge of revealing the whole sentence, one word at a  
204 time, using a keyboard in front of them. They could either unmask forward or backward (to  
205 reveal again words they already read) as many times as they wanted. When participants  
206 considered they had finished reading the sentence accurately (no matter which word was  
207 unmasked at that moment), they said the word “stop” and the experimenter stopped the trial.  
208 A preliminary training phase with short sentences (French proverbs) was performed to  
209 familiarize the participant with the task and protocol. When the participant was comfortable  
210 with the keys and instructions, the test phase began and the blocks of sentences were  
211 presented.

212 Reading time (in seconds) was programmatically recorded for each word as the time during  
213 which it was unmasked. For words unmasked only once, total reading time was equal to the  
214 first-pass reading time (i.e. the first – *and only* - unmasked period). For words unmasked  
215 several times, total reading time was the sum of all the unmasked periods.



216

217 **Figure 2: Self-paced Reading paradigm**

218 Words are revealed one at a time by the participant, using keyboard presses. At any given time  
 219 during a trial, a press on the right arrow key will unmask the upcoming word (while masking  
 220 the currently visible word) and a press on the left arrow key will unmask the previous word  
 221 (while masking the currently visible word). For words unmasked only once (e.g. word 1), the  
 222 total reading time equals the 1<sup>st</sup>-pass reading time. For words unmasked more than once (e.g.  
 223 word 2), the total reading time is the sum of each individual reading time pass. Target word T is  
 224 located at the end of the sentence but is not necessarily unmasked at the end of the trial, as in  
 225 this example where the participant unmasked the word T-1 to confirm it had been read  
 226 accurately before responding to end the trial.

227

228 **2.5 Statistical analysis**

229 Statistical analyses were carried out using R (Team, 2018) with the following additional R-  
 230 packages: ggplot2, tidyr, dplyr, gridextra, lme4, stringr. Analyses 1 and 2 were performed  
 231 with linear mixed-effects models that allow the modeling of experimental designs with

232 unbalanced repeated measures (Bolker et al., 2009; Cheng et al., 2010). For each analysis,  
233 several models were first constructed with total reading time and accuracy of each target word  
234 as dependent variables. Reading time and word frequency value were transformed in natural  
235 logarithm (ln) units. The continuous variables were centered around their mean. In a second  
236 step, the Akaike Information Criterion (AIC) and likelihood-ratio tests were used to assess an  
237 optimal random-effects structure (Zuur et al., 2010). Once an optimal model structure had  
238 been chosen, significance of the fixed effects was estimated using t-values. Usually, the  
239 number of degrees of freedom for the t-values of the fixed effects are not exactly known with  
240 mixed-effects models (Baayen et al., 2008; Pinheiro & Bates, 2000). However, given the large  
241 number of observations in the present study, the t-distribution converges to a normal  
242 distribution. Therefore, following standard recommendations, t-values (i.e. estimate/standard  
243 error) that were larger than 2 were considered significant, which corresponds to a significance  
244 level of 5% in a two-tailed test (Baayen et al., 2008; Gelman & Hill, 2007). Finally,  
245 assumptions underlying the models were visually checked with diagnostic plots of residuals  
246 (Cohen et al., 2003; Pinheiro & Bates, 2000). In the Results section fixed-effects estimates are  
247 reported along with their t-values and 95% confidence intervals (Bates et al., 2015).

248

249

### 250 **3. RESULTS**

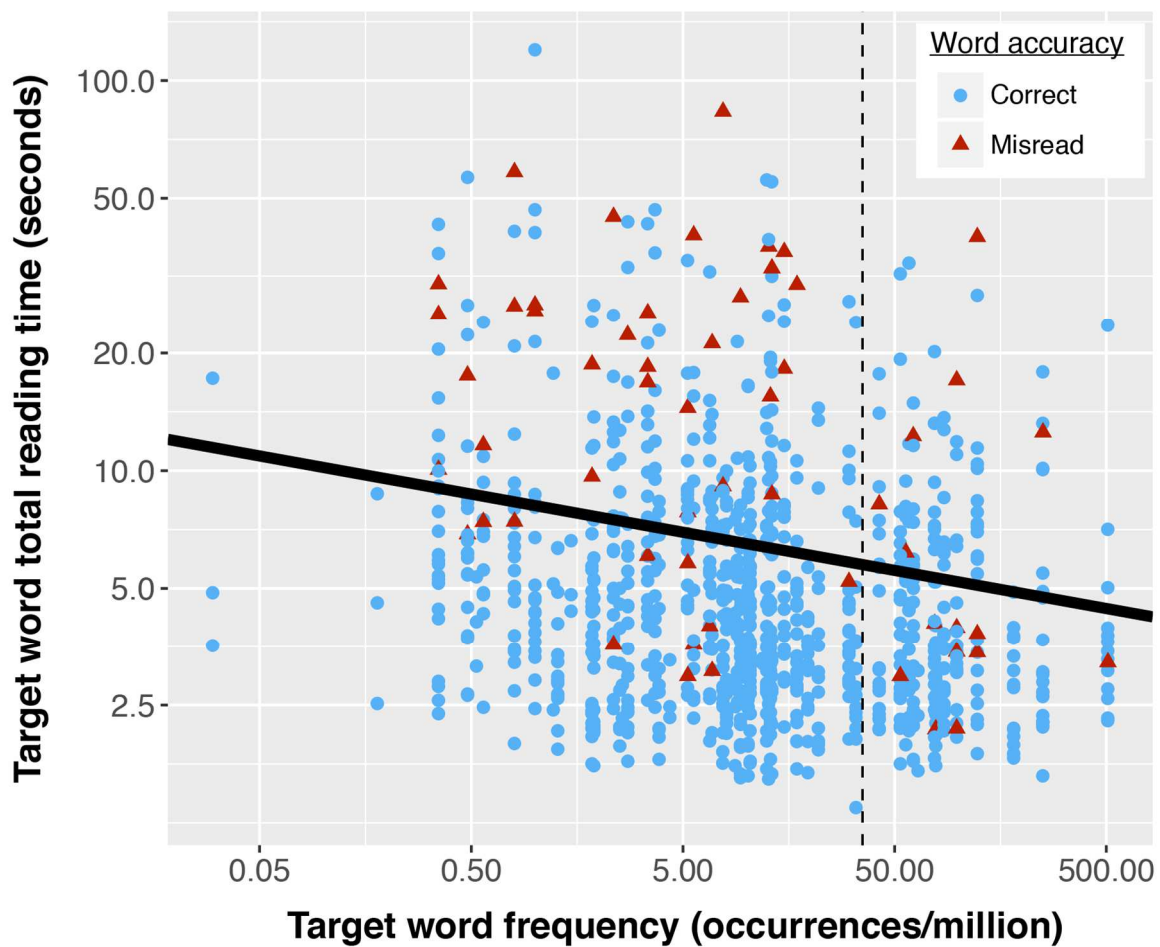
251

#### 252 **3.1 Analysis 1: Effect of word frequency value on reading performance**

253 Here we tested the effect of frequency value of the target words (in occurrences/million) on  
254 both reading accuracy and total reading time. On average, target words were read accurately  
255 93% of the time, with individual variations ranging from 69 to 100% depending on  
256 participants. We found that the effect of word frequency on reading accuracy was not

257 significant (estimate = 0.005, 95% CI = [-0.003; 0.01]), probably due to ceiling effects. Figure  
258 3 shows the target word total reading time of as a function of target word frequency, for all  
259 trials and participants. This relationship was further analysed for correct trials only with a  
260 linear mixed-effects model containing the following factors as fixed effects: target word  
261 frequency, word length (in number of characters), age, low-vision rehabilitation history (yes  
262 or no) and “still reading” (yes or no). The random effects were the following: subjects, per  
263 subject frequency slopes, target words and condition (nested within target words). The mixed-  
264 effects analysis showed a significant effect of word frequency on reading time, as represented  
265 by the black solid line in Figure 3. Results of the model are shown in Table 2.

266



267

268 **Figure 3: Scatterplot of target word total reading time as a function of word frequency value**

269 **(Analysis 1).** Data are plotted for target words read both correctly (blue circles) and incorrectly



270 (red triangles). The black solid line represents the slope of the frequency effect for words read  
 271 correctly only, as estimated by the mixed-effects model. The dashed vertical line represents the  
 272 mean word frequency value in initial units (about 35 occurrences/million). Note the log scales  
 273 used for both axes.

274

	Estimate	Standard error	t-value	95% Confidence Interval
Intercept	2.09	0.025	8.3	[1.6; 2.56]
<b>Word frequency (ln)</b>	<b>-0.10</b>	<b>0.02</b>	<b>-4.98</b>	<b>[-0.138; -0.06]</b>
Word length (n. of characters)	-0.00004	0.023	-0.042	[-0.046; 0.044]
Age (years)	0.003	0.006	0.551	[-0.008; 0.015]
Low-vision rehabilitation <i>No</i>	-0.063	0.197	-0.317	[-0.454; 0.321]
<b>Still reading</b> <i>No</i>	<b>-0.558</b>	<b>0.225</b>	<b>-2.482</b>	<b>[-0.984; -0.132]</b>

275

276 **Table 2: Results of the mixed-effects model from Analysis 1: fixed effects estimates.**

277 **The dependent variable is target word total reading time. The reference level of categorical**  
 278 **binary factors is italicized. Factors showing a significant effect on reading time are in bold font.**

279

280 The main factor of interest, word frequency, had a significant effect with a regression  
 281 coefficient estimate of -0.10 (95% CI = [-0.138; -0.06]). As both reading time and frequency  
 282 are expressed in natural log units, this means that multiplying frequency (in original units) by  
 283 10 multiplies reading time (in original units) by 0.79 ( $10^{-0.10}$ ), i.e. a 21 % decrease.  
 284 Similarly, we can also express the effect's amplitude in the following way: multiplying  
 285 frequency (in original units) by 1000 (from 0.5 to 500, i.e. the range where most of our values  
 286 lie) multiplies reading time (in original units) by 0.5 ( $1000^{-0.10}$ ), i.e. a 50 % decrease. It is  
 287 notable that low-frequency words are on average longer than high-frequency words, so that  
 288 the frequency effect could have been induced by the confound between word length and word  
 289 frequency (Kliegl et al., 1982; Rayner & Duffy, 1986) – the correlation was -0.17 in the  
 290 present study. This possibility was however eliminated here by partialling out the effect of

291 word length in the model. Finally, we did not find any interaction between word frequency  
292 and word length (estimate = -0.002, 95% CI = [-0.02; 0.02]). This result is not consistent with  
293 previous studies of natural reading in normally sighted skilled readers based on gaze duration  
294 (Hyönä & Olson, 1995; Kliegl et al., 2006). Because age is a predictor of reading speed, even  
295 for people with normal visual acuity and self-reported healthy vision (Calabrèse et al., 2016;  
296 Lott et al., 2001; Warrington et al., 2018), we controlled for its effect by including this factor  
297 in our model. Although age was found to be non significant, it must be controlled for because  
298 it correlates with other factors, such as motor or attention deficits, that influence reading  
299 speed (Sass et al., 2006). The last factor with a statistically significant effect on reading time  
300 was the binary “still reading” (estimate = -0.558, 95% CI = [-0.984; -0.132]). According to  
301 the model, participants who still read, even a few minutes each day, are expected to have their  
302 word reading time multiplied by 0.57 ( $\exp[-0.558]$ ), i.e. a decrease of 43%. Note that similar  
303 results were found when using first-pass reading time as the dependent variable, instead of  
304 total reading time (Figure 2).

305

### 306 **3.2 Analysis 2: Effect of word frequency category on reading performance**

307 Within each pair of synonyms, we tested the effect of frequency category (low vs. high) of the  
308 target words on both reading accuracy and total reading time. First we found that frequency  
309 category did not have a significant effect on target word accuracy within pairs (estimate = -  
310 0.009, 95% CI = [-0.04; 0.02]): low-frequency synonyms were not read less accurately than  
311 their higher frequency counterpart. To analyse the effect of frequency category on target word  
312 total reading time, we removed all pairs containing one or two inaccurately read words.  
313 The linear mixed-effects model was the same as in Analysis 1 except for two points. First,  
314 frequency category (low vs. high) was used as a fixed effect (instead of the continuous  
315 frequency value factor). Second, the random structure was different to accommodate for the

316 within-pair analysis so that the optimal model contained: the subject factor, the pair factor and  
 317 the by-pair category frequency factor as random effects. The results of the model are shown in  
 318 Table 3.

319

	Estimate	Standard error	t-value	95% Confidence interval
intercept	2.21	0.27	8.2	[1.7; 2.7]
<b>Frequency category</b> <i>Low frequency</i>	<b>-0.133</b>	<b>0.046</b>	<b>-2.882</b>	<b>[-0.227; -0.043]</b>
Word length (n. of characters)	0.051	0.034	1.498	[-0.016; 0.118]
Age (years)	0.003	0.007	0.501	[-0.009; 0.016]
Low vision rehabilitation <i>No</i>	-0.188	0.214	-0.88	[-0.588; 0.211]
<b>Still reading</b> <i>No</i>	<b>-0.505</b>	<b>0.24</b>	<b>-2.104</b>	<b>[-0.953; -0.057]</b>

320

321 **Table 3: Results of the mixed-effects model from Analysis 2: fixed effects estimates.**

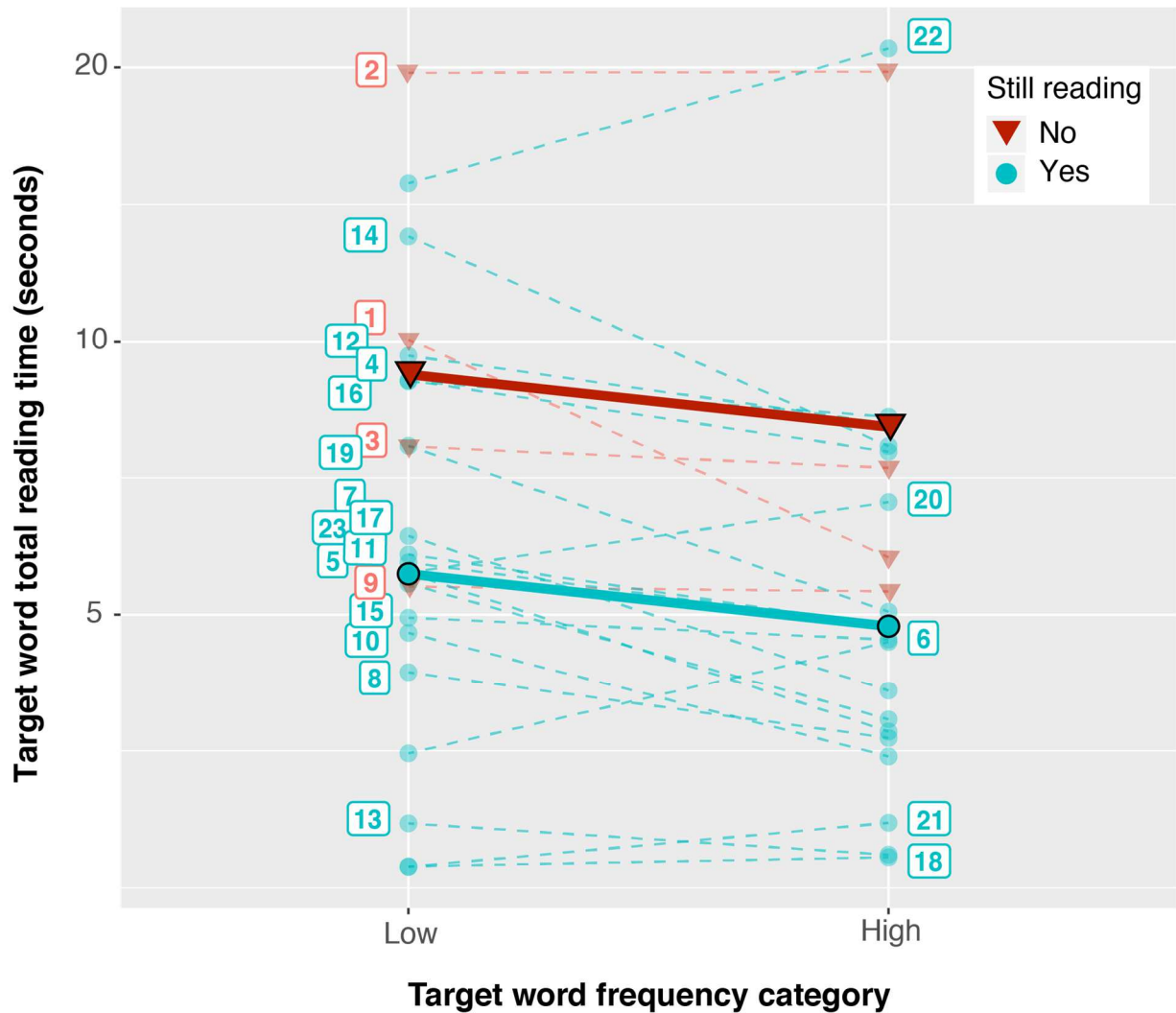
322 **The dependent variable is the target words total reading time. The reference level of categorical**  
 323 **binary factors is italicized. Factors with a significant effect on reading time are in bold font.**

324

325 Within-pair frequency category was found to be a significant predictor of target word total  
 326 reading time (estimate = -0.13; 95% CI=[-0.23; -0.04]). This means that high-frequency  
 327 words have a reading time that is multiplied by 0.87 ( $\exp[-0.133]$ ) compared to their low-  
 328 frequency synonyms (i.e. a 13 % decrease). This effect is represented in Figure 4 by thick  
 329 solid lines. The dashed lines in this figure show the mean values for each participant,  
 330 emphasizing that the amplitude of the frequency effect at the population level (solid lines) is  
 331 quite small compared to the between-subjects differences. Despite being statistically  
 332 significant, the 13% decrease in reading time induced by higher frequencies does not seem  
 333 impressive in terms of clinical relevance. This relatively modest decrease is most likely due to  
 334 the limited frequency ratio assigned to synonyms of each pair. Indeed, the within-pair  
 335 frequency ratio was set between 5 and 10 in the present work (Figure 1 - step1), with an

336 average of 7.07. Using results from Analysis 1 to predict the amplitude of the frequency  
 337 category effect without by-pair grouping, we find that total reading time for high-frequency  
 338 words would be multiplied by 0.82 ( $7.07^{-0.10}$  – see Table 2) compared to their low-  
 339 frequency synonyms. This 18% decrease, predicted from Analysis 1, is thus very close to the  
 340 13% reduction found in Analysis 2. We hypothesize that increasing the ratio of frequencies  
 341 within pair of synonyms should boost the frequency category effect up to a clinically relevant  
 342 improvement. Finally, we found that the “still reading” factor had a significant effect on total  
 343 reading time (Figure 4 & Table 3), similar to the one reported in Analysis 1.

344  
 345



346  
 347

348 **Figure 4: Effect of frequency category on target word reading time for both the “still reading”**  
349 **(blue circles) and the “not still reading” (red inverted triangles) groups (Analysis 2). Solid lines**  
350 **connect the estimates for each group as given by the mixed-effects model. Dashed lines connect**  
351 **the mean reading time values for each participant, labeled from 1 to 23. Individual trends**  
352 **(downward vs. upward) are represented by the labels’ position (left vs. right). Note the log scale**  
353 **for the y-axis.**

354

355

356

357

358

#### 359 **4. DISCUSSION**

360

361 The general objective of this work was to investigate whether and how the degraded visual  
362 input of individuals with central field loss alters the processes linking low-level visual  
363 analysis of words to lexical access. More precisely, our first purpose was to assess the effect  
364 of word frequency on word reading time when individuals with bilateral maculopathy are  
365 engaged in natural reading. The dependent variable in our study was the reading time of target  
366 words (included within sentences) whose visual and linguistic properties are accurately  
367 controlled. This has never been studied before in visually impaired patients during natural  
368 reading (i.e. eye-mediated reading).

369

370 We used the method of self-paced reading as it allowed us to measure the reading time of  
371 individual words either for first-pass reading or for second- and nth-pass reading. Thus, when  
372 a word is unmasked for long periods (e.g. 10 sec), we can safely assume that many fixations  
373 have been made and constitute a cluster as previously described (Calabrèse et al., 2016).

374 Results obtained with this method can be fruitfully compared to those from eye movement  
375 studies of natural reading (Clifton Jr. et al., 2016; Pollatsek et al., 2008; Rayner, 1998). In this  
376 huge literature, one important dependent variable is "gaze duration", classically defined as the  
377 total fixation time on word  $n$  before gaze position moves on to word  $n+1$  for the first time:  
378 this might correspond to one or several fixations (when several fixations are made, they are  
379 called "refixations" on the same word). Another measure is "total fixation duration" (sum of  
380 all fixations on a word  $n$ , including those induced by regressive saccades from words to the  
381 right of word  $n$ ). Gaze duration is typically interpreted as reflecting earlier processes than  
382 those measured by total fixation duration. In our study, we assume that our two dependent  
383 variables, first-pass reading time and total reading time, are relevant proxies of gaze duration  
384 and total fixation duration (the only approximation being that durations of saccades are  
385 ignored). It is worth noting that in normal reading, short connector words and articles are  
386 often skipped entirely by gaze position, whereas the self-paced reading paradigm forces  
387 subjects to fixate on each word of the sentence, possibly modifying their reading behavior.  
388 However, this point is not relevant in the present study because none of the target words were  
389 short connector words or articles.

390

391 Our results show a clear-cut effect of word frequency on word reading time (whether first-  
392 pass reading time or total reading time) even after partialling out the effect of word length  
393 (Kliegl et al., 1982; Rayner & Duffy, 1986). This result is consistent with recent findings  
394 reporting that reading speed was significantly reduced by the presence of clusters of fixations  
395 and that these clusters seemed located near the words with the lowest frequency (Calabrèse et  
396 al., 2016). This result is also in line with the effect of word frequency for subjects with normal  
397 vision: higher frequency words yield faster reading speed. However, the effect of word  
398 frequency found in our CFL participants is much greater than what has been reported before

399 for normally sighted subjects. In an influential natural reading study with skilled readers, it  
400 was reported that gaze duration decreases from 240 to 207 ms when word frequency is  
401 multiplied by 1000 – from 1 to 4 log 10 units (Kliegl et al., 2006). Despite some differences in  
402 methodology between this work and the present study, the 14 % decrease in gaze duration  
403 reported in Kliegl et al., 2006 can be compared to the much larger 50 % decrease that we  
404 report in Analysis 1 for the same range of word frequencies. Such difference suggests that  
405 low-vision individuals rely more on lexical inference than normally sighted subjects.

406

407 Our work might be related to a previous study investigating the effect of word frequency on  
408 lexical decision accuracy with normally sighted observers (Lee et al., 2003). The task was to  
409 decide whether stimuli of different durations and seen either with foveal or peripheral vision  
410 were words or pseudo-words. The authors reported that significant frequency effects occurred  
411 for the shortest durations in central vision, 25–50 ms, whereas significant frequency effects  
412 did not occur in peripheral vision until 100 ms. Although it is difficult to predict how these  
413 effects relate to a natural reading task, this study was important because it raised the major  
414 question of lexical access in peripheral vision. This important issue was also addressed in “Mr  
415 Chips”, an implemented ideal-observer model of low-vision reading (Legge et al., 2002;  
416 Legge et al., 1997). Because this model integrates visual and lexical information to predict the  
417 pattern of eye fixations required to read a given text, lexical access constraint is expected to  
418 induce frequency effects on fixation duration.

419

420 Our results do not show any evidence for an effect of word length on word reading time  
421 although this effect is commonly reported in eye movement research on natural reading: word  
422 reading time increases with word length, mainly as a result of the increasing number of  
423 “refixations” (Hyönä & Olson, 1995; Kliegl et al., 2004; Kliegl et al., 2006; Kliegl et al.,

424 1982; Vitu et al., 2001; Vitu et al., 1990). Typically, normally sighted readers would make  
425 one single fixation on 5-character words and two fixations on 10-character words. This  
426 apparent discrepancy about the word length effect between CFL and normally sighted readers  
427 is probably due to diverse factors. First, the range of word lengths used in our study is  
428 probably too small (relative to the high variability in the dependent variable) to uncover an  
429 effect of word length whose amplitude and robustness are known to be smaller than the  
430 frequency effect for instance. Word length goes from 5 to 10 characters in the present study,  
431 whereas the range of lengths can be as high as 2 to 20 in some studies (Kliegl et al., 2006).  
432 Second, it is possible that the effect of word length is absent, or difficult to measure, in our  
433 study because people with CFL systematically make many fixations on nearly all the words of  
434 a sentence (Calabrèse et al., 2016; Calabrèse et al., 2014; Scherlen et al., 2008). This high  
435 base level would make any subtle difference such as the word length effect more difficult to  
436 extract.

437

438 The secondary purpose of our work was to consider the potential of text simplification as a  
439 valuable tool allowing to improve reading performance of visually impaired patients. One of  
440 the many possibilities offered by text simplification is to replace complex words by their  
441 simpler synonyms (Saggion, 2017), a transformation which requires agreement on a definition  
442 of word complexity. In this work, we decided to use word frequency as a surrogate of word  
443 complexity. Our second statistical analysis was thus performed on pairs of target word  
444 synonyms and showed an effect of relative frequency category (low vs. high) on word reading  
445 time: participants read target words of higher relative frequency faster than their lower  
446 frequency synonyms.

447



448 Importantly, the amplitude of this frequency category effect could be approximated from the  
449 mixed-effects model run in Analysis 1, which did not take into account the by-pair grouping.  
450 This result is encouraging because it shows that a moderate frequency effect can be observed  
451 between synonyms having a relatively moderate difference in frequency (the frequency ratio  
452 ranged from 5 to 10 in the present study, with an average of 7) over a very large range of  
453 frequencies. In other words, our results do not apply only to the low frequency words of the  
454 French language. They apply to a large range of frequencies (say between 0.5 and 50  
455 occurrences per million - Figure 3) and thus concern a very large number of words. As long as  
456 the numerous words lying within this range have synonyms of slightly higher frequency (ratio  
457 > about 7), a benefit in reading time is predicted by our statistical analysis. Of course, if  
458 synonyms with a higher ratio than 10 exist, they should be chosen preferentially by text  
459 simplification algorithms as the predicted reading time decrease would be larger.

460

461 Recent advances in the domain of natural language processing, at the frontier between  
462 linguistics computer science and artificial intelligence, should allow a large-scale  
463 implementation of text simplification methods. Among the current possibilities offered by text  
464 simplification, lexical simplification based on word frequency, among other factors, is easy to  
465 implement in automated algorithms and could be applied to various contexts (in electronic  
466 books, web sites, etc.). Hopefully, reducing the complexity of lexical units in text, without  
467 changing meaning, should improve overall reading performance of low-vision readers. We  
468 feel that the most important aspect of this improvement would not be the increased  
469 performance (e.g. in reading speed) *per se* but the possibility of keeping a certain level of  
470 motivation for reading activities. What our results clearly show is that people who keep on  
471 reading are those who read faster. This correlation between the motivation to read and word  
472 reading speed may create either a vicious circle or a virtuous one. It seems therefore important

473 to help low-vision persons fight against this pattern. Our hope is that text simplification might  
474 encourage low-vision readers to read more and more often, which would, in turn, maintain a  
475 certain level of practice preventing them from stopping reading activities altogether.

476

477

## 478 **CONCLUSION**

479

480 This study showed a robust and large amplitude effect (+ 50%) of word frequency on word  
481 reading time when participants with macular diseases read text. The strength of this effect  
482 suggests that lexical access in CFL readers heavily relies on inferential processes compared to  
483 normally sighted observers. Our study also highlighted the potential benefits of lexical  
484 simplification when replacing a word by a synonym of higher frequency. More generally, the  
485 long-term challenge of this work is to investigate what aspects of visual, lexical and sentence  
486 processing can be simplified in order to significantly improve reading performance of low-  
487 vision patients.

488

489 One asset of our study is that the amplitude of the frequency effect is quantified by a  
490 regression coefficient while controlling for other relevant factors. This should allow future  
491 research to quantitatively compare whether the frequency effect is more capable of improving  
492 reading speed than other factors. Similarly, regression analyses should also help decide if the  
493 effects of frequency or other factors are the same for different categories of reading deficits  
494 (e.g. dyslexia, aphasia). If differential effects were found, this would suggest that methods of  
495 text simplification should be adapted to the specific needs of different reading disorders.

496

497

498 **ACKNOWLEDGMENTS**

499

500 This work was supported by the Fondation de France and by the Belgian FNRS. The authors

501 thank Dr Frédéric Chouraqui and Dr Alban Comet for their help in recruiting the participants.

502 **REFERENCES**

503

504 Aaronson D, Scarborough HS (1976) Performance theories for sentence coding: Some  
505 quantitative evidence. *Journal of Experimental Psychology: Human Perception and*  
506 *Performance* 2: 56-70.

507

508 Baayen RH, Davidson DJ, Bates DM (2008) Mixed-effects modeling with crossed random  
509 effects for subjects and items. *Journal of Memory and Language* 59: 390-412.

510

511 Bates D, Mächler M, Bolker B, Walker S (2015) Fitting Linear Mixed-Effects Models Using  
512 lme4. *Journal of Statistical Software* 67: 1-48.

513

514 Bernard JB, Scherlen AC, Castet E (2007) Page mode reading with simulated scotomas: a  
515 modest effect of interline spacing on reading speed. *Vision Research* 47: 3447-59.

516

517 Bernard J, Aguilar C, Castet E (2016) A New Font, Specifically Designed for Peripheral  
518 Vision, Improves Peripheral Letter and Word Recognition, but Not Eye-Mediated Reading  
519 Performance. *PLoS One* 11: e0152506.

520

521 Billami M, François T, Gala N (2018) ReSyf: a French lexicon with ranked synonyms. In  
522 *Proceedings of COLING 2018, August, 20-26, Santa Fe, USA.*

523 <https://cental.uclouvain.be/resyf/>.

524

525 Blackmore-Wright S, Georgeson MA, Anderson SJ (2013) Enhanced text spacing improves  
526 reading performance in individuals with macular disease. *PLoS One* 8: e80325.

527

528 Bolker BM, Brooks ME, Clark CJ, Geange SW, Poulsen JR, Stevens MH, et al. (2009)  
529 Generalized linear mixed models: a practical guide for ecology and evolution. *Trends Ecol*  
530 *Evol* 24 (3): 127-135.

531

532 Bowers AR, Woods RL, Peli E (2004) Preferred retinal locus and reading rate with four  
533 dynamic text presentation formats. *Optom Vis Sci* 81: 205-13.

534

535 Brown JC, Goldstein JE, Chan TL, Massof R, Ramulu P, Low Vision Research Network  
536 Study Group (2014) Characterizing functional complaints in patients seeking outpatient low-  
537 vision services in the United States. *Ophthalmology* 121: 1655-62.e1.

538

539 Calabrèse A, Bernard J, Faure G, Hoffart L, Castet E (2016) Clustering of Eye Fixations: A  
540 New Oculomotor Determinant of Reading Speed in Maculopathy. *Invest Ophthalmol Vis Sci*  
541 57: 3192-202.

542

543 Calabrèse A, Bernard J, Faure G, Hoffart L, Castet E (2014) Eye movements and reading  
544 speed in macular disease: the shrinking perceptual span hypothesis requires and is supported  
545 by a mediation analysis. *Invest Ophthalmol Vis Sci* 55: 3638-45.

546

547 Calabrèse A, Bernard J, Hoffart L, Faure G, Barouch F, Conrath J, et al. (2010) Small effect  
548 of interline spacing on maximal reading speed in low-vision patients with central field loss  
549 irrespective of scotoma size. *Invest Ophthalmol Vis Sci* 51: 1247-54.

550

551 Calabrèse A, Cheong AMY, Cheung S, He Y, Kwon M, Mansfield JS, et al. (2016) Baseline

552 MNREAD Measures for Normally Sighted Subjects From Childhood to Old Age. Invest  
553 Ophthalmol Vis Sci 57: 3836-43.  
554

555 Calabrèse A, Liu T, Legge GE (2017) Does Vertical Reading Help People with Macular  
556 Degeneration: An Exploratory Study. PLoS One 12: e0170743.  
557

558 Cheng J, Edwards LJ, Maldonado-Molina MM, Komro KA, Muller KE (2010) Real  
559 longitudinal data analysis for real people: building a good enough mixed model. Stat Med 29:  
560 504-20.  
561

562 Chung STL (2002) The effect of letter spacing on reading speed in central and peripheral  
563 vision. Invest Ophthalmol Vis Sci 43: 1270-6.  
564

565 Chung STL, Jarvis SH, Woo SY, Hanson K, Jose RT (2008) Reading speed does not benefit  
566 from increased line spacing in AMD patients. Optom Vis Sci 85: 827-33.  
567

568 Chung STL, Legge GE, Cheung SH (2004) Letter-recognition and reading speed in peripheral  
569 vision benefit from perceptual learning. Vision Research 44: 695-709.  
570

571 Chung STL, Mansfield JS, Legge GE (1998) Psychophysics of reading XVIII - The effect of  
572 print size on reading speed in normal peripheral vision. Vision Research 38: 2949-62.  
573

574 Chung STL (2011) Improving reading speed for people with central vision loss through  
575 perceptual learning. Invest Ophthalmol Vis Sci 52: 1164-70.

576 Clifton Jr. C, Ferreira F, Henderson JM, Inhoff AW, Liversedge SP, Reichle ED, et al. (2016)  
577 Eye movements in reading and information processing: Keith Rayner's 40 year legacy. *J.*  
578 *Mem. Lang.* 86: 1-19.  
579  
580 Cohen J, Cohen P, West SG, Aiken LS (2003) *Applied multiple regression/correlation*  
581 *analysis for the behavioral sciences.* Mahwah, NJ: Lawrence Erlbaum Associates.  
582  
583 Elliott DB, Trukolo-Ilic M, Strong JG, Pace R, Plotkin A, Bevers P (1997) Demographic  
584 characteristics of the vision-disabled elderly. *Invest Ophthalmol Vis Sci* 38: 2566-75.  
585  
586 Fine EM, Peli E (1998) Benefits of rapid serial visual presentation (RSVP) over scrolled text  
587 vary with letter size. *Optom Vis Sci* 75: 191-6.  
588  
589 Gelman A, Hill J (2007) *Data analysis using regression and multilevel/hierarchical models.*  
590 Cambridge University Press, New York.  
591  
592 Hyönä J, Olson RK (1995) Eye fixation patterns among dyslexic and normal readers: effects  
593 of word length and word frequency. *J Exp Psychol Learn Mem Cogn* 21: 1430-40.  
594  
595 Just MA, Carpenter PA, Woolley JD (1982) Paradigms and processes in reading  
596 comprehension. *J Exp Psychol Gen* 111: 228-38.  
597  
598 Klein R, Klein BE, Cruickshanks KJ (1999) The prevalence of age-related maculopathy by  
599 geographic region and ethnicity. *Prog Retin Eye Res* 18: 371-89.  
600

601 Kliegl R, Grabner E, Rolfs M, Engbert R (2004) Length, frequency, and predictability effects  
602 of words on eye movements in reading. *European Journal of Cognitive Psychology* 16: 262-  
603 284.

604

605 Kliegl R, Nuthmann A, Engbert R (2006) Tracking the mind during reading: the influence of  
606 past, present, and future words on fixation durations. *J Exp Psychol Gen* 135: 12-35.

607

608 Kliegl R, Olson RK, Davidson BJ (1982) Regression analyses as a tool for studying reading  
609 processes: comment on Just and Carpenter's eye fixation theory. *Mem Cognit* 10: 287-96.

610

611 Kumar A, Chawla R, Kumawat D, Pillay G (2017) Insight into high myopia and the macula.  
612 *Indian J Ophthalmol* 65: 85-91.

613

614 Lee HW, Legge GE, Ortiz A (2003) Is word recognition different in central and peripheral  
615 vision? *Vision Res* 43: 2837-46.

616

617 Legge GE (2007) *Psychophysics of reading in normal and low vision*. Mahwah: NJ &  
618 London: Lawrence Erlbaum Associates.

619

620 Legge GE, Ahn SJ, Klitz TS, Luebker A (1997b) Psychophysics of reading--XVI. The visual  
621 span in normal and low vision. *Vision Res* 37: 1999-2010.

622

623 Legge GE, Hooven TA, Klitz TS, Mansfield JS, Tjan BS (2002) Mr. Chips 2002: new  
624 insights from an ideal-observer model of reading. *Vision Res* 42: 2219-234.

625



626 Legge GE, Klitz TS, Tjan BS (1997a) Mr. Chips: an ideal-observer model of reading. Psychol  
627 Rev 104: 524-53.  
628

629 Li C, Lin K, Lin Y, Lee J (2002) Low vision and methods of rehabilitation: a comparison  
630 between the past and present. Chang Gung Med J 25: 153-61.  
631

632 Lott LA, Schneck ME, Haegerström-Portnoy G, Brabyn JA, Gildengorin GL, West CG  
633 (2001) Reading performance in older adults with good acuity. Optom Vis Sci 78: 316-24.  
634

635 Mangione CM, Berry S, Spritzer K, Janz NK, Klein R, Owsley C, et al. (1998) Identifying the  
636 content area for the 51-item National Eye Institute Visual Function Questionnaire: results  
637 from focus groups with visually impaired persons. Arch Ophthalmol 116: 227-33.  
638

639 Mansfield JS, Legge GE, Bane MC (1996) Psychophysics of reading XV - Font effects in  
640 normal and low vision. Invest Ophthalmol Vis Sci 37: 1492-501.  
641

642 Meyniel C, Bodaghi B, Robert P (2017) Revisiting Vision Rehabilitation. Front Syst Neurosci  
643 11: 82.  
644

645 Mitchell DC, Green DW (1978) The effects of context and content on immediate processing  
646 in reading. Q. J. Exp. Psychol. 30: 609-636.  
647

648 New B, Brysbaert M, Veronis J, Pallier C (2007) The use of film subtitles to estimate word  
649 frequencies. Applied Psycholinguistics 28: 661-677.  
650

651 Peirce JW (2007) PsychoPy--Psychophysics software in Python. *J Neurosci Methods* 162: 8-  
652 13.  
653

654 Peirce JW (2009) Generating stimuli for neuroscience using PsychoPy. *Frontiers in*  
655 *Neuroinformatics* 2: 10.  
656

657 Pelli DG (2008) Crowding: a cortical constraint on object recognition. *Curr Opin Neurobiol*  
658 18: 445-51.  
659

660 Pinheiro JC, Bates DM (2000) *Mixed-effects models in S and S-Plus*. Springer, .  
661

662 Pollatsek A, Juhasz BJ, Reichle ED, Machacek D, Rayner K (2008) Immediate and delayed  
663 effects of word frequency and word length on eye movements in reading: a reversed delayed  
664 effect of word length. *Journal of Experimental Psychology: Human Perception and*  
665 *Performance* 34: 726--750.  
666

667 R Core Team (2018) *R: A Language and Environment for Statistical Computing*. Vienna,  
668 Austria: R Foundation for Statistical Computing. , <https://www.R-project.org/>.  
669

670 Rayner K (1998) Eye movements in reading and information processing: 20 years of research.  
671 *Psychol Bull* 124: 372-422.  
672

673 Rayner K, Duffy SA (1986) Lexical complexity and fixation times in reading: effects of word  
674 frequency, verb complexity, and lexical ambiguity. *Mem Cognit* 14: 191-201.  
675

676 Rubin GS, Turano K (1994) Low vision reading with sequential word presentation. *Vision*  
677 *Res* 34: 1723-33.  
678

679 Rubin GS, Turano K (1992) Reading without saccadic eye movements. *Vision Res* 32: 895-  
680 902.  
681

682 Saggion H (2017) Automatic Text Simplification. *Synthesis Lectures on Human Language*  
683 *Technologies* 10: 1-137.  
684

685 Sass SM, Legge GE, Lee HW (2006) Low-vision reading speed: influences of linguistic  
686 inference and aging. *Optom Vis Sci* 83: 166-77.  
687

688 Scherlen AC, Bernard JB, Calabrese A, Castet E (2008) Page mode reading with simulated  
689 scotomas: oculo-motor patterns. *Vision Res* 48: 1870-8.  
690

691 Shardlow M (2014) A survey of automated text simplification. *Int. J. Adv. Comput. Sci.*  
692 *Appl. Special Issue on Natural Language Processing 2014*: doi:10.14569.  
693

694 Siddharthan A (2014) A survey of research on text simplification. *Int. J. Appl. Linguist.* 165:  
695 259-298.  
696

697 Silva R (2012) Myopic maculopathy: a review. *Ophthalmologica* 228: 197-213.  
698

699 Steen-Baker AA, Ng S, Payne BR, Anderson CJ, Federmeier KD, Stine-Morrow EAL (2017)  
700 The effects of context on processing words during sentence reading among adults varying in

701 age and literacy skill. *Psychol Aging* 32: 460-472.

702

703 Vitu F, McConkie GW, Kerr P, O'Regan JK (2001) Fixation location effects on fixation  
704 durations during reading: an inverted optimal viewing position effect. *Vision Res* 41: 3513-  
705 33.

706

707 Vitu F, O'Regan JK, Mittau M (1990) Optimal landing position in reading isolated words and  
708 continuous text. *Percept Psychophys* 47: 583-600.

709

710 Vongphanit J, Mitchell P, Jin Wang J (2002) Prevalence and progression of myopic  
711 retinopathy in an older population. *Ophthalmology* 109: 704-11.

712

713 Warrington KL, McGowan VA, Paterson KB, White SJ (2018) Effects of aging, word  
714 frequency, and text stimulus quality on reading across the adult lifespan: Evidence from eye  
715 movements. *J Exp Psychol Learn Mem Cogn.*

716

717 Wolffsohn JS, Cochrane AL (1999) The changing face of the visually impaired: the Kooyong  
718 low vision clinic's past, present, and future. *Optom Vis Sci* 76: 747-54.

719

720 Wong WL, Su X, Li X, Cheung CMG, Klein R, Cheng C, et al. (2014) Global prevalence of  
721 age-related macular degeneration and disease burden projection for 2020 and 2040: a  
722 systematic review and meta-analysis. *Lancet Glob Health* 2: e106-16.

723

724 Yu D, Cheung SH, Legge GE, Chung STL (2007) Effect of letter spacing on visual span and  
725 reading speed. *J Vis* 7: 7.

726

727 Zuur AF, Ieno EN, Elphick CS (2010) A protocol for data exploration to avoid common

728 statistical problems. *Methods in Ecology and Evolution* 1: 3-14.

729

730

731 **APPENDIX 1**

732

733 C'est en rentrant dans la cour que j'ai vu mon **garçon (fiston)**

734 J'aimerais bien aller à la maison avec mon **fiston (garçon)**

735 Il est dommage qu'il n'ait pas pu montrer ses **capacités (aptitudes)**

736 Il aurait vraiment pu essayer de développer ses **aptitudes (capacités)**

737 La petite fille demande si elle peut faire des **biscuits (galettes)**

738 Il a proposé de venir chez elle et d'apporter des **galettes (biscuits)**

739 Après tout ce qu'il lui a dit, elle ne peut que le **détester (mépriser)**

740 Plus le temps passait, plus il était amené à la **mépriser (détester)**

741 Pour bien dessiner, tout est question de bonnes **habitudes (pratiques)**

742 Pour faire ce métier, il faut avoir de bonnes **pratiques (habitudes)**

743 Cette décision n'a rien à voir avec une **punition (sanction)**

744 Je pense qu'il aurait préféré une autre **sanction (punition)**

745 Cet homme grand a quelque chose en lui de très **étrange (suspect)**

746 Il est entré dans la maison et il avait l'air **suspect (étrange)**

747 Après toutes ces années, voici sa plus belle **découverte (trouvaille)**

748 Tout le monde a été impressionné par cette **trouvaille (découverte)**

749 Je sais qu'il a toujours de l'air chaud qui sort de sa **narine (naseau)**

750 Il utilise une solution spéciale pour nettoyer chaque **naseau (narine)**

751 C'est sur les grands boulevards qu'a eu lieu cette **marche (défilé)**

752 Ils se sont préparés longtemps pour organiser ce **défilé (marche)**

753 Il reste de la terre car ce trou était facile à **remplir (combler)**

754 Ses enfants l'ont vécu comme un vide difficile à **combler (remplir)**

755 Toute son équipe savait qu'il a donné sa vie à sa **recherche (invention)**

- 756 Ils veulent faire connaître au monde entier leur **invention (recherche)**
- 757 Je lui ai montré une peinture avec un magnifique **buisson (arbuste)**
- 758 Si tu vas derrière la pharmacie, tu verras un étrange **arbuste (buisson)**
- 759 En politique, comme en amour, on a parfois besoin de **repos (trêve)**
- 760 Etant donnée ton attitude, tu ne mérites pas de **trêve (repos)**
- 761 Ils ont essayé par tous les moyens de le **corrompre (débaucher)**
- 762 Ce n'est pas faute d'avoir tenté de les **débaucher (corrompre)**
- 763 L'homme est bon par nature, la société ne fait que le **dégrader (dépraver)**
- 764 Avec son passé, cela n'a pas été très difficile de le **dépraver (dégrader)**
- 765 Cette association compte plus de curieux que de **disciples (adhérents)**
- 766 Il faudrait encore leur demander d'appeler les **adhérents (disciples)**
- 767 Elle m'a prévenu qu'elle voulait qu'on respecte son **souhait (vouloir)**
- 768 Cependant, il n'est pas question qu'on suive son **vouloir (souhait)**
- 769 Pour réussir ce plat, il faut choisir une belle **salade (laitue)**
- 770 Si tu y penses, rapporte de la lessive et une **laitue (salade)**
- 771 Il aurait fallu être très malin pour le **prévoir (prédire)**
- 772 Il faut que tu arrêtes d'essayer de tout **prédire (prévoir)**
- 773 Le passé composé est un temps très facile à **utiliser (employer)**
- 774 Tu disposes de nombreux outils, il faut les **employer (utiliser)**
- 775 Je compte sur toi pour m'aider à trouver la **somme (total)**
- 776 Il a remarqué que cela ne correspond pas au **total (somme)**
- 777 Pendant le rendez-vous, elle l'a trouvé complètement **dingue (cinglé)**
- 778 L'autre jour il leur est arrivé quelque chose de **cinglé (dingue)**
- 779 Les fêtes approchent et tu ne dois pas oublier de le **rendre (livrer)**
- 780 Ils l'ont prévenu qu'il n'avait que deux jours pour se **livrer (rendre)**

- 781 Pour réparer cet ordinateur, il faut faire une nouvelle **demande (requête)**
- 782 Il a pris un rendez-vous et son patron pourra recevoir sa **requête (demande)**
- 783 Il était gêné de demander une culotte de femme au **commerçant (colporteur)**
- 784 Je n'ai pas assez d'argent pour aller chez ce **colporteur (commerçant)**
- 785 Je dois deviner où tu as rangé ce **dossier (archive)**
- 786 Je n'ai pas réussi à retrouver cette **archive (dossier)**
- 787 J'ai rencontré un homme avec un humour **piquant (mordant)**
- 788 Ses commentaires manquent souvent de **mordant (piquant)**
- 789 Penser est pour lui un accident au lieu d'un état **permanent (continuel)**
- 790 Il faut rester calme et serein même si c'est un défi **continuel (permanent)**
- 791 Cette nouvelle m'a laissé dans un état de grand **abattement (sidération)**
- 792 Il a traversé une phase de colère puis de profonde **sidération (abattement)**
- 793 C'est pour cela que toute la famille part faire une **marche (balade)**
- 794 Il avait besoin de changer d'air et est allé faire une **balade (marche)**
- 795 Après avoir travaillé sur cette loi, il a décidé de la **retarder (ajourner)**
- 796 Ce n'est pas possible de perdre encore du temps et de l'**ajourner (retarder)**
- 797