Special Issue Article

QJEP

Quarterly Journal of Experimental Psychology
2018, Vol. 71(1) 20–27

DOI: 10.1080/17470218.2016.1237531
qjep@sagepub.com
SAGE

Eye movements when reading sentences with handwritten words

Manuel Perea, Ana Marcet, Beatriz Uixera and Marta Vergara-Martínez

Abstract

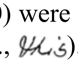
The examination of how we read handwritten words (i.e., the original form of writing) has typically been disregarded in the literature on reading. Previous research using word recognition tasks has shown that lexical effects (e.g., the word-frequency effect) are magnified when reading difficult handwritten words. To examine this issue in a more ecological scenario, we registered the participants' eye movements when reading handwritten sentences that varied in the degree of legibility (i.e., sentences composed of words in easy vs. difficult handwritten style). For comparison purposes, we included a condition with printed sentences. Results showed a larger reading cost for sentences with difficult handwritten words than for sentences with easy handwritten words, which in turn showed a reading cost relative to the sentences with printed words. Critically, the effect of word frequency was greater for difficult handwritten words than for easy handwritten words or printed words in the total times on a target word, but not on first-fixation durations or gaze durations. We examine the implications of these findings for models of eye movement control in reading.

Keywords
Eye movements; reading; word recognition

Received: 28 April 2016; accepted: 5 September 2016

Despite being the original form of writing, the examination of how we process handwritten words has been overlooked in the literature on reading (see Rayner, Pollatsek, Ashby, & Clifton, 2012, for a review). As the visual input is inherently noisy and variable, reading handwritten text presents a number of challenges (see Manso De Zuniga, Humphreys, & Evett, 1991): (a) The shape of letters not only differs in each instance but can be affected by the neighbouring letters; and (b) some of the words' component letters may be connected, thus leading to an extra segmentation process. While there has been some recent interest on this issue in the area of visual-word recognition (e.g., Barnhart & Goldinger, 2010, 2013, 2015; Gil-López, Perea, Moret-Tatay, & Carreiras, 2011; Perea, Gil-López, Beléndez, & Carreiras, 2016; Qiao et al., 2010), this has not yet been translated into more ecological, sentence reading experiments comparing handwritten and printed sentences when the individuals' eye movements are registered.

Previous evidence using isolated word recognition tasks (e.g., lexical decision) has revealed that top-down processes may exert a greater role on handwritten words than on printed words. Leaving aside that, unsurprisingly, word identification times are longer for handwritten words than for typed words, Manso De Zuniga et al. (1991) and Barnhart and Goldinger (2010, 2013) found that the word-frequency effect (i.e., an indicator of lexical processing) was greater for handwritten words than for printed words in the lexical decision task—Barnhart and Goldinger (2010, 2015) also found a magnification of other lexical effects with handwritten words (e.g., imageability, consistency, and orthographic neighbourhood). To explain these findings, Barnhart and Goldinger (2010) indicated that when reading handwritten words, the human perceptual system "simply has to rely more heavily on top-down processes, relative to more prototypical word forms"(p. 921).

The handwritten words employed by Barnhart and Goldinger (2010) were "highly non-uniform and unfamiliar" (p. 908; e.g., ). Notably, Barnhart and Goldinger (2010) indicated that, in the original design of their

Departamento de Metodologia and ERI-Lectura, Universitat de Valencia, Valencia, Spain

Corresponding author:
Manuel Perea, Departamento de Metodologia and ERI-Lectura, Universitat de Valencia, Av. Blasco Ibanez, 21, Valencia 46010, Spain.
Email: mperea@uv.es

experiments, they had included a human print condition that yielded the same pattern of results as printed words—however, no specific details were provided. To inspect in detail the role of the legibility of handwritten words during lexical access, Perea et al. (2016) conducted a series of lexical decision experiments that included two handwritten styles that varied in their legibility. Specifically, they employed words written by an individual with poor penmanship (difficult handwritten words; e.g., música, puñal— note the variable letter shape, spacing, and alignment) and words written by an individual with good penmanship (easy handwritten words; e.g., música, puñal —note the uniform letter shape and alignment). To select the two individuals with good versus poor penmanship, Perea et al. (2016) asked eight individuals to write down 10 sentences, and 10 naïve raters assessed the readability of the sentences —Perea et al. (2016) chose the individuals with the best and worst penmanship (mean scores of 4.2 and 2.3 in a 1-to-5 scale, respectively). Words printed in Century font (e.g., música, puñal) were used as a control. Perea et al. (2016, Experiment 3) found faster word identification times for printed words (573 ms) than for easy handwritten words (587 ms), which in turn were identified faster than difficult handwritten words (619 ms). More important, the magnitude of the word-frequency effect was greater for difficult handwritten words (63 ms) than for easy handwritten words (46 ms) or printed words (44 ms). To explain this pattern of findings, Perea et al. (2016) argued that, early in processing, handwritten words involve some reading cost at mapping visual features to letters (see Manso De Zuniga et al., 1991, for discussion). This would readily explain the additive effects of script and word frequency for easy handwritten words. In addition, the difficult handwritten words would require additional top-down processing— note that, unlike easy handwritten words, the processing of difficult handwritten words involves additional activation in frontoparietal brain areas (see Qiao et al., 2010, for functional magnetic resonance imaging, fMRI, evidence). The idea is that top-down processes exert a greater influence when the bottom-up information is not strong, thus increasing lexical effects (e.g., the word-frequency effect; see Barnhart & Goldinger, 2010, 2015). That is, the larger word-frequency effect for difficult handwritten words can be interpreted as a "balance of bottom-up and top-down activity" (Perea et al., 2016, p. 1644) in interactive models of visual-word recognition (see Carreiras, Armstrong, Perea, & Frost, 2014, for a recent review).

Although lexical decision experiments are certainly informative and valuable, they only offer a single data point at the end of processing (i.e., the word identification time). Instead, eye movement experiments during sentence reading allow for a much richer analysis of the time course of the effects (e.g., early vs. late eye movement measures). Furthermore, sentence reading experiments offer a more ecological scenario than presenting isolated words. In the current experiment, we employed sentences composed of printed words, easy handwritten words, and difficult handwritten words that contained a target word of high or low frequency (see Table 1). This allows us to examine not only global sentence reading measures (e.g., total time, fixation duration, number of progressive/regressive saccades), but also local measures on the target word. Keep in mind that lexical decision times cannot easily disentangle whether the enhanced word frequency for difficult words is due to early encoding, lexical access, or post-lexical integration mechanisms (see Gomez, Perea, & Ratcliff, 2013, for an attempt to dissociate encoding vs. decision processes in lexical decision with the diffusion model). To examine this question in the context of sentence reading, it may be important to consider one of the leading models of eye movement control in reading, the E-Z Reader model (Reichle, Pollatsek, Fisher, & Rayner, 1998;see also Reichle & Sheridan, 2015, for a recent review of the model). After an initial pre-attentive stage of visual word processing that extracts low-level visual information from the retina, lexical access in the E-Z Reader model takes places in two stages: an initial (and fast) sense of familiarity check (L1 stage) and a slower retrieval of semantic information that also involves completion of lexical access (L2 stage). Specifically, L1 begins when allocation of attention is located on the target word and ends when the word is about to be identified. When L1 is completed, a saccade is programmed to the next word. As L1 is modulated by word frequency, the E-Z Reader model can readily explain why there are fewer and shorter fixations for high- than for low-frequency words. The E-Z Reader model can accommodate the presence of regressions to the target word by assuming that there was a failure at a post-lexical integration stage (see Reichle, Warren, & McConnell, 2009)—note that regressions can also occur if the saccade to the next word occurs while attention is still directed to the previous word (due to extended L2). This would explain why some effects do not occur when measuring early first-pass measures (e.g., the first-fixation duration on the target word, or in the sum of fixations on the target word before leaving it—gaze duration), but they are noticeable when measuring the total times on the target

Table 1. Depiction of the three conditions in the experiment.

Sentence type	Sentence
Difficult handwritten sentences	Mi amigo es inglés aunque reside en España
Easy handwritten sentences	Mi amigo es inglés aunque reside en España
Printed sentences	Mi amigo es inglés aunque reside en España
English translation	[My friend is English although he lives in Spain]

Note: In the experiment, handwitten and printed sentences occupied the same horizontal space.

word (e.g., neighbourhood frequency effect; see Perea & Pollatsek, 1998;Slattery, 2009). Other models of eye movement control in reading make similar predictions in this respect (e.g., SWIFT model, Engbert, Nuthmann, Richter, & Kliegl, 2005). Given that most phenomena initially found in word identification tasks have been generalized to sentence reading (e.g., neighbourhood frequency, letter transposition, etc.; see Perea & Pollatsek, 1998; White, Johnson, Liversedge, & Rayner, 2008), one would expect a main effect of script (i.e., a reading cost for handwritten sentences when compared to printed sentences). Importantly, at the local level, the magnification of the word-frequency effect with difficult handwritten words would have a different interpretation depending on whether it occurs in early, first-pass measures (e.g., probability of first-pass fixation, first-fixation duration, gaze duration) or in later measures that reflect post-lexical processing stages (e.g., total time).

To our knowledge, no previous experiments have compared handwritten vs. printed words during sentence reading when the individuals' eye movements are registered. Nonetheless, a number of experiments have examined how fonts of differing legibility affect eye movement control during reading (see Slattery, 2016, for a recent review on reading and font design). Rayner, Reichle, Stroud, Williams, and Pollatsek (2006) employed sentences in Times New Roman (i.e., a standard an easily legible font) or Old English (i.e., an uncommon and difficult font) that contained a high- or low-frequency word. At the global sentence level, Rayner et al. (2006) found a substantial cost of reading sentences in Old English relative to Times New Roman (i.e., more fixations, longer fixation durations, and shorter saccades) in adult skilled readers. Importantly, at the local level, the size of the word-frequency effect on the target word was similar for Old English and Times New Roman in first-pass eye movement measures (first-fixation durations: 20 vs. 16 ms; gaze durations: 56 vs. 48 ms, respectively). Indeed, it was only in the total times on the target word that the magnitude of the word-frequency effect was greater in Old English than in Times New Roman (88 vs. 55 ms, respectively). In another study, Slattery and Rayner (2010, Experiment 1) compared three fonts: Times New Roman, Harrington (i.e., an uncommon and hard-to-read font), and Script MT (i.e., an uncommon font in cursive). At the global level, Slattery and Rayner reported a sizeable reading cost of the Harrington and Script MT fonts when compared to Times New Roman (i.e., more fixations, longer total reading times, longer fixation durations). At the local level, Slattery and Rayner found a word-frequency effect of comparable magnitude across fonts for first-fixation durations (word-frequency effect: 22, 8, and 28 ms for Times New Roman, Harrington, and Script MT, respectively) and gaze durations (word-frequency effect: 35, 67, and 49 ms in Times New Roman, Harrington, and Script MT, respectively), and it was only in the total times that they found a substantially greater word-frequency effect for the Harrington and Script MT fonts (128 and 122 ms, respectively) than for Times New Roman (35 ms).

Taken together, the effect of font legibility in the Rayner et al. (2006) and Slattery and Rayner (2010) experiments only interacted significantly with word frequency in a late measure of eye movements (i.e., total times), but not in earlier first-pass eye movement measures (i.e., first-fixation durations, gaze durations). As Slattery and Rayner (2010) indicated: "it is likely that these variables are affecting a post access processing stage rather than affecting initial access" (p. 1145). What is the nature of this alleged post-access stage? Slattery and Rayner suggested two possibilities: (a) Some of the words could have been initially misidentified (Slattery, 2009); and (b) readers "maintain uncertain beliefs about the identities of previously read words" (Levy, Bicknell, Slattery, & Rayner, 2009, p. 21089), so that some of the words may need to be re-inspected. An additional assumption is that these processes should be affected by both script legibility and word frequency (Slattery & Rayner, 2010).

To sum up, the present experiment examined sentence reading for printed versus (easy/difficult) handwritten words. Previous research using isolated word recognition in lexical decision showed greater lexical effects (e.g., word frequency) with difficult handwritten words than for easy handwritten or printed words. If the magnification of word-frequency effect for difficult handwritten words during sentence reading takes place in the initial access to these words, it should be observable in first-pass measures (first-fixation durations, gaze durations). Alternatively, if the magnification of the word-frequency effect for difficult handwritten words during sentence reading is due to late word identification processes or to post-lexical difficulty/failure arising from integration with other words in the sentence (i.e., syntax, discourse representation), it would be reflected in the total times on the target word.

Experimental study

Method

Participants. Twenty-four students from the University of Valencia, all of them native speakers of Spanish, participated in the experiment in exchange of a small gift. All participants had normal/corrected-to-normal vision.

Materials. The 120 sentence frames were the same as those in the Perea and Acha (2009) sentence reading experiment. Each sentence frame included a target word that could be of high or low frequency [e.g., "Mi amigo es inglés/griego aunque reside en España" (My friend is English/Greek although he lives in Spain)]. As Perea and Acha showed, the sentence frames with high- and low-frequency words were equally understandable, and the target word was not predictable from the context in a cloze task. We created two sets of materials composed of 60 sentences with a high-frequency target word [average word frequency per million = 87; average length = 7.3, range = 5–9; average number of orthographic neighbours (Coltheart's *N*) = 0.58, in the B-Pal database; Davis & Perea, 2005] and 60 sentences with a low-frequency target word (mean word frequency per million = 4.5; average length = 7.3, range = 5–9; average number of orthographic neighbours = 0.96). The number of orthographic neighbours was similar in each set ($p > .50$). Sentences were presented typewritten, written from someone with a good penmanship, or written from someone with a poor penmanship—all the handwritten sentences were scanned. None of the sentences included spelling errors. To create the handwritten sentences, in a pilot stage of the experiment, we asked 10 individuals to write six sentences. Eight students who did not know the purpose of the experiment rated the level of penmanship of each individual on a 1 – 7 Likert scale. We chose the individuals with the best and worst penmanship (average score of 6.4 vs. 2.6, respectively) to write the sentences for the experiment (see Perea et al., 2016, for a similar procedure; see Table 1). These two individuals were asked to write down the entire sentences as they would normally do, and to monitor that each word in the sentence had a similar length to those in printed script (14-pt Times New Roman)—this was done by having a sheet with the printed script just below the sheet used by the individuals to write down the sentences. For each participant, 40 sentences were presented in printed script (20 with an embedded high-frequency target word, 20 with an embedded low-frequency target word), 40 sentences with easy handwritten words, and 40 sentences with difficult handwritten words—we created three counterbalanced lists in each set.

Apparatus. To record the individuals' eye movements, we employed a camera-based Eyelink 1000 K eyetracker (SR Research Ltd, Canada). The sample pupil location was 1000 Hz. Viewing was binocular, but the device only recorded data from the right eye. Sentences were presented on a single line of a CRT monitor (22" ViewSonic Professional series P225f) linked to a Windows-OS computer.

Procedure. The participant was seated approximately 60 cm from the monitor in a quiet, faintly lit room. A chinrest was used to reduce head motion. We employed Eyetrack software (http://www.psych.umass.edu/eyelab/software/) to present the sentences/questions and to collect the eye fixation data. In each trial, the participant was asked to look at a black square in the left part of the screen. Once the participant gazed at the black square, a single-line sentence appeared—the initial letter of the sentence corresponded to the black square. Participants were asked to read each sentence for comprehension, as they would normally do—they were also asked to refrain from blinking (if possible) when reading the sentences. They were also instructed to press a button from a gamepad to terminate the trial. They were also told that, after a quarter of the sentences, they would be presented with a yes/no comprehension question on the sentence they had just read by pressing the button "si" (yes) or "no" from the gamepad. Before the experimental phase, there was a 9-point calibration and validation phase. This was followed by nine practice sentences. Calibration was checked at the start of every trial—recalibration was performed when required. The order of the sentences was randomized for each participant.

Data analysis. The EyeDoctor and Eyedry software suites (http://www.psych.umass.edu/eyelab/software/) were used to process the eye movement data. In an initial step, the very short fixations (<80 ms) that were within one letter of the previous/next fixation were merged. In a second step, single fixations beyond the 80–800-ms cut-offs (less than 2% data) were excluded from the eye fixation analyses. We conducted F_1 (by-subject) and F_2 (by-item) analyses of variance (ANOVAs). For the local analyses, the factors were: (a) word frequency of the target word (low, high) and (b) script (printed, easy handwritten, difficult handwritten). We examined the following dependent variables: probability of first-pass fixation on the target word, first-fixation duration on the target word in the first pass, gaze duration (i.e., sum of first-pass fixations on the target word before leaving it), and total time on the target word. Finally, to depict the expected interaction between script and word frequency for total time, we also examined the go-past time (i.e., the sum of fixations on from first fixating the target word until moving to the next word, thus including re-reading earlier words), percentage of regressions back to target word.

Results

The accuracy rate in the yes/no comprehension questions was over 90% in all script conditions (printed sentences: 96.3%; easy handwritten sentences: 95.4%; difficult handwritten sentences 91.6%)—Bonferroni t-tests only revealed a marginal (non-significant) advantage of the printed condition over the difficult handwritten condition ($p = .066$). The average data for each experimental condition are presented in Tables 2 and 3. Although the main focus of the present experiment is on the measures at the level of the target word (i.e., local analyses), we now briefly present the analyses on several relevant dependent variables at the global sentence level: total reading time, fixation duration, and number of fixations (both progressive and regressive).

Global analyses. The ANOVAs showed a dramatic effect of script (all Fs>50, all ps<.001; see Table 2 for details): Total reading times and fixation durations were longer for difficult

Table 2. Global measures for total sentence reading time, mean fixation duration, and number of saccades.

Script	Total reading time	Mean fixation duration	Number of saccades	
			Progressive	Regressive
Printed	2471 (212)	236 (7.2)	7.7 (0.5)	2.7 (0.3)
Handwritten (easy)	2925 (234)	247 (7.2)	8.8 (0.5)	3.0 (0.3)
Handwritten (difficult)	4552 (390)	275 (8.5)	11.2 (0.7)	4.9 (0.5)
Easy handwritten-printed	*ps* < .001	*ps* < .001	*ps* < .001	p_1 = .016 p_2 < .001
Difficult-easy handwritten	*ps* < .001	*ps* < .001	*ps* < .001	*ps* < .001

Note: Sentence reading time in ms; fixation duration in ms. Standard errors in parentheses.

Table 3. Local measures on the target word for each of the conditions: percentage of first-pass fixations on the target word, first-fixation duration, gaze duration, go-past time, percentage of regressions, and total time.

Script	Fix (%)		FFD (ms)		GD (ms)		Go-past time (ms)		Regr. in (%)		TT (ms)	
	LF	HF	LF	HF	LF	HF	LF	HF	LF	HF	LF	HF
Printed	93.9	90.8	249	234	351	293	405	340	14.5	12.7	449	376
	(1.3)	(1.1)	(8)	(8)	(26)	(17)	(44)	(35)	(2.3)	(1.7)	(42)	(31)
Easy handwritten	96.8	95.4	273	257	430	373	530	445	13.8	10.4	562	471
	(1.1)	(1.0)	(11)	(7)	(34)	(20)	(55)	(30)	(2.3)	(2.0)	(51)	(31)
Difficult handwritten	98.3	97.1	299	277	554	463	771	562	23.4	19.1	903	682
	(0.7)	(0.9)	(11)	(9)	(35)	(32)	(67)	(40)	(2.1)	(2.5)	(66)	(63)

Note: Fix = fixations on the target word; FFD = first-fixation duration; GD = gaze duration; Regr. in = regressions; TT = total time. LF = low-frequency words; HF = high-frequency words. Standard errors in parentheses.

handwritten sentences than for easy handwritten sentences (4552 vs. 2925 ms and 275 vs. 247 ms, respectively), which in turn were longer than the reading times and fixation durations for printed sentences (2462 and 236 ms). Likewise, the number of progressive saccades was higher for difficult handwritten sentences than for easy handwritten sentences (11.2 vs. 8.8, respectively) and, in turn, for printed sentences (7.7). Finally, we found the same pattern in the number of regressive saccades: It was higher for difficult handwritten sentences than for easy handwritten sentences (4.9 vs. 3.0, both *ps* <.001) and printed sentences (2.7, both *ps* <.0165).

Local analyses

Probability of first-pass fixation on the target word. The probability of fixating the target word was higher for low- than for high-frequency words (96.4 vs. 94.4%, respectively) [$F_1(1,23) = 8.05$, *MSE* = 16.4, p =.009; $F_2(1, 118) =$ 4.58, *MSE* = 74.3, *p* =.034]. The main effect of script was also significant [$F_1(2, 46) = 14.17$, *MSE* = 25.4, p <.001; $F_2(2, 236) = 16.58$, *MSE* =53.9, p <.001; printed words (92.4%) < easy handwritten words (96.1%) < difficult handwritten words (97.7%); all *ps* <.045]. There were no signs of an interaction between the two factors (both *Fs* < 1).

First-fixation durations. First-fixation durations on the target word were shorter for high- than for low-frequency words (256 vs. 273 ms, respectively) [$F_1(1, 23) = 20.9$, *MSE* = 541.6, p <.001; $F_2(1, 118) = 13.47$, *MSE* = 1702.2, *p* <.001]. The main effect of script was also significant [$F_1(2, 46) = 40.57$, *MSE* =636.4, *p* <.001; $F_2(2, 236) =$ 44.57, *MSE* = 611.1, *p*<.001; typed words (242 ms) < easy handwritten words (265 ms) < difficult handwritten words (288 ms), all *ps* <.001]. As can be seen in Table 3, the two effects were additive (interaction: both *Fs* < 1).

Gaze durations. Gaze durations on the target word were shorter for high- than for low-frequency words (376 vs. 445 ms, respectively) [$F_1(1, 23) = 41.43$, *MSE* = 4090.0, *p* <.001; $F_2(1, 118) = 27.17$, *MSE* =13,213.3, *p* <.001]. The main effect of script was also significant [$F_1(2, 46)$ = 93.62, *MSE* =4479.0, *p* <.001; $F_2(2, 236) = 99.55$, *MSE* =10,244.4, *p* <.001; typed words (322 ms) < easy handwritten words (401 ms) < difficult handwritten words (508 ms), all ps <.001]. Although the magnitude of the word-frequency effect was numerically larger for the difficult handwritten words (91 ms) than for the easy handwritten words (57 ms) or the typed words (58 ms), the interaction between the two factors did not approach significance [$F_1(2, 46) = 1.13$, *MSE* = 3965, *p* =.33, $F_2(2, 236) = 1.11$, *MSE* = 10,244.4, *p* =.33].

Total times. Total times on the target word were shorter for high- than for low-frequency words (510 vs. 638 ms, respectively) [$F_1(1, 23) = 59.17$, *MSE* = 10,052.1, *p*<.001;

$F_2(1, 118) = 25.80$, $MSE= 56{,}476.8$, $p <.001$]. The main effect of script was also significant [F_1 (2, 46) = 85.27, $MSE = 21{,}683.7$, $p <.001$; $F_2(2, 236) = 150.50$, $MSE = 56{,}476.8$, $p <.001$; typed words (412 ms) < easy handwritten words (517 ms) < difficult handwritten words (792 ms), all *ps* <.001]. The interaction between the two factors was significant [$F_1(2, 46) = 7.86$, $MSE = 9868.3$, $p =.001$; $F_2(2, 236) = 6.56$, $MSE = 30{,}215.4$, $p=.002$]. This interaction showed that the magnitude of the word-frequency effect was dramatically larger for difficult handwritten words (221 ms) than for easy handwritten words (91 ms) or printed words (73 ms); all *ps* <.001.

Go-past times. Go-past times on the target word were shorter for high- than for low-frequency words (449 vs. 569 ms, respectively) [$F_1(1,23) = 26.79$, $MSE = 19{,}292.3$, $p <.001$; $F_2(1, 118) = 37.54$, $MSE = 38{,}212.4$, $p <.001$]. The main effect of script was also significant [$F_1(2, 46) = 115.85$, $MSE =9091.3$, $p <.001$; $F_2(2, 236) = 113.94$, $MSE= 26{,}070.9$, $p<.001$; typed words (373 ms) < easy handwritten words (487 ms) < difficult handwritten words (666 ms), all *ps* <.001]. The interaction between the two factors was significant [$F_1(2, 46) = 15.12$, $MSE = 4911.3$, $p <.001$; $F_2(2, 236) = 10.24$, $MSE = 26{,}070.9$, $p <.001$]; the magnitude of the word-frequency effect was substantially larger for difficult handwritten words (210 ms) than for easy handwritten words (85 ms) or printed words (64 ms); all *ps* <.001.

Percentage of regressions back to target word. The percentage of regressions back to the target word was lower for high- than for low-frequency words (14.0 vs. 17.2%, respectively) in the analysis by subjects [$F_1(1, 23) = 4.97$, $MSE = 72.9$, $p =.036$; $F_2(1, 118) = 3.46$, $MSE = 242.7$, $p =.065$]. The main effect of script was significant [F_1 (2, 46) = 17.92, $MSE = 64.7$, $p <.001$; $F_2(2, 236) = 13.48$, MSE = 206.1, p <.001; easy handwritten words (12.1%) = printed words (13.6% ms) < difficult handwritten words (21.2%), all *ps*<.001]. Finally, there were no signs of an interaction between the two factors (both *Fs* < 1).

Discussion

In the current experiment, we compared sentence reading with printed words versus easy/difficult handwritten words. At the global sentence level, results showed a considerable reading cost of handwritten sentences when compared to printed sentences (e.g., longer sentence reading times, longer fixation durations, more fixations—both progressive and regressive). As expected, this difference was modulated by the legibility of the handwritten style (i.e., a greater reading cost for the difficult handwritten sentences than for the easy handwritten sentences). As each sentence contained a low- or high-frequency word, we also examined the interplay between script and word frequency at the local level—note that prior lexical decision experiments reported a magnification of the word-frequency effect for difficult, but not for easy, handwritten words (Barnhart & Goldinger, 2010; Perea et al., 2016). In early first-pass eye movement measures, we found additive effects of script and word frequency: (a) the probability of fixating the target word was higher for low-frequency words than for high-frequency words, and it was higher for handwritten words than for printed words (difficult handwritten > easy handwritten > printed); and (b) first-fixation durations and gaze durations were longer for low-frequency words than for high-frequency words, and they were longer for handwritten words than for printed words (difficult handwritten > easy handwritten > printed). The magnification of the word-frequency effect occurred in a late measure—namely, the total time on the target word: The word-frequency effect was considerably larger for difficult handwritten words than for easy handwritten words and printed words.

To characterize the interaction between script and word frequency for total time, we examined two measures that could help reveal the time course of this effect. First, we measured the percentage of regressions back to the target word. This measure revealed a higher percentage of regressions back to the target word for the difficult handwritten sentences. The absence of an interaction for the regression in measure suggests that the interaction in total time is not driven by post-lexical integration failure driving regressions in. Nonetheless, we need to take into account that re-reading times are longer for low- than for high-frequency words (Raney & Rayner, 1995), and, hence, this measure could have modulated the total time for low- and high-frequency words. Second, we examined the go-past time (i.e., the sum of all fixations from first fixating the target word until fixation the following word/s to the right). In this case, we found a robust interaction between script and word frequency: The magnitude of the word-frequency effect in go-past times was substantially larger for difficult handwritten words than for easy handwritten words or printed words—we also examined the first-pass regressions out, and the critical interaction approached significance in the by-subjects analyses, $F_1(2, 46) = 2.44$, p =.098 (the word-frequency effect was slightly higher for difficult handwritten words than for the easy handwritten words). As only go-past and total time measures show a robust interaction between script and word frequency, then the larger word-frequency effect for difficult handwritten words could arise due to the cumulative effect of re-reading. That is, the interaction might be driven by late word identification processes ("post-access processing stage"; see Slattery & Rayner, 2010), not necessarily integration failure. Nonetheless, as suggested by a reviewer, it is also possible that the go-past and total times could reflect some integration difficulty with other words in the sentence (e.g., arising due to misidentifications).

This pattern of data closely resembles that obtained with uncommon fonts that are difficult to read. As indicated in the introduction, Rayner et al. (2006) and Slattery and Rayner (2010) found, at the global sentence level, a substantial overall reading cost for difficult unfamiliar fonts (Old English font in the Rayner et al., 2006, experiment; Harrington font in the Slattery & Rayner, 2010, study). Importantly, as in the Rayner et al. (2006) and the Slattery and Rayner (2010) experiments, this was not translated into a magnification of the word-frequency effect in first-pass eye movement measures at the local level (e.g., probability of fixating the target word, first-fixation duration, gaze duration). These findings suggest that script (i.e., printed vs. handwritten words in the current experiment) does not modulate, at least not to a large degree, the initial stage of lexical processing. Instead, the magnification of the word-frequency effect primarily occurred in those measures that include re-reading. In other words, the greater word-frequency effect for difficult handwritten words than for easy handwritten words or printed words occurred because these words were re-inspected.[1] As put forward by Slattery and Rayner (2010), this could occur because the target word was initially mis-identified (Slattery, 2009) or because the level of uncertainty regarding the identity of the current/previous words reaches a threshold and requires re-inspection (Levy et al., 2009; see also Bicknell & Levy, 2010)—the present experiment was not aimed to test these two options.

Thus, the present data suggest that there may be an initial encoding cost at mapping letter features to abstract letter representations for all handwritten words. This would be reflected in longer fixation times and more saccades for easy and difficult handwritten than for printed sentences, and it may occur at a pre-attentive (pre-lexical) processing stage in the framework of the E-Z Reader model. In addition, difficult handwritten words yielded an additional reading cost at a post-access processing stage, possibly when integrating the visual percept with the (prototypical) stored representations in lexical/semantic memory—note that this may involve a greater cost for lower than for higher frequency words (see Figure 3 in Reichle, 2015, for a depiction of how failures in post-lexical integration processes produce regressive saccades in the E-Z Reader model).

We acknowledge that a broad range of factors may potentially modulate the pattern of effects in handwritten sentences. As a reviewer pointed out, handwriting can differ across a range of different variables such as joined versus separated letters, visual complexity, crowding, and difference between the individual letter format compared to the standard, among others. Future research should carefully examine which of these factors may be critical to interactions with lexical or sublexical factors.

To conclude, the present experiment represents an initial step to examine how handwriting affects sentence reading. We have shown that, while handwritten sentences yield an important reading cost when compared to printed sentences (i.e., more and longer fixations than printed text), they do not seem to greatly affect the initial stages of lexical processing—as deduced from the additive effects of script (printed vs. handwritten) and word frequency in first-pass eye movement measures. Clearly, the inherent variability of handwritten text presents a series of challenges for models of eye movement control in reading, and further research should focus on how higher level processes (e.g., predictability) are affected by the noisy visual input from handwritten sentences. As Manso De Zuniga et al. (1991) claimed, "our understanding of word recognition, and, by extension of reading, will at best be incomplete if it is confined to typescript" (p. 11).

Acknowledgements

We thank Steve Goldinger and two anonymous reviewers for very helpful comments on earlier versions of this paper.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

The research reported in this article has been partially supported by the Spanish Ministry of Economy and Competitiveness [grant number PSI2014-53444-P], [grant number BES-2015-07414].

ORCiD

Manuel Perea http://orcid.org/0000-0002-3291-1365
Ana Marcet http://orcid.org/0000-0001-8755-5903
Marta Vergara-Martínez http://orcid.org/0000-0003-1931-0335

Note

1. We acknowledge that, in the present data, as also occurs in the data from Rayner et al. (2006) and Slattery and Rayner (2010), there are some hints that suggest that the word-frequency effect can be slightly greater for the "difficult" condition in the gaze durations. Nonetheless, the bulk of the magnification of the word-frequency effect occurs in the total times on the target word.

Supplementary Material

Supplemental material is available at journals.sagepub.com/doi/suppl/10.1080/17470218.2016.1237531.

References

Barnhart, A. S., & Goldinger, S. D. (2010). Interpreting chicken-scratch: Lexical access for handwritten words. *Journal of Experimental Psychology: Human Perception and Performance*, *36*, 906–923. doi:10.1037/a0019258

Barnhart, A. S., & Goldinger, S. D. (2013). Rotation reveals the importance of configural cues in handwritten word perception. *Psychonomic Bulletin and Review*, *20*, 1319–1326. doi:10. 3758/s13423-013-0435-y

Barnhart, A. S., & Goldinger, S. D. (2015). Orthographic and phonological neighborhood effects in handwritten word perception. *Psychonomic Bulletin and Review*, *22*, 1739–1745. doi:10. 3758/s13423-015-0846-z

Bicknell, K., & Levy, R. (2010). A rational model of eye movement control in reading. In J. Hajic, S. Carberry, S. Clark & J. Nivre (Eds.), Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics (ACL (pp. 1168–1178)). Uppsala: Association for Computational Linguistics.

Carreiras, M., Armstrong, B. C., Perea, M., & Frost, R. (2014). The what, when, where, and how of visual word recognition. *Trends in Cognitive Sciences*, *18*,90–98. doi:10.1016/j. tics. 2013.11.005

Davis, C. J., & Perea, M. (2005). BuscaPalabras: A program for deriving orthographic and phonological neighborhood statistics and other psycholinguistic indices in Spanish. *Behovior Research Methods*, *37*, 665–671. doi:10.3758/ BF03192738

Engbert, R., Nuthmann, A., Richter, E. M., & Kliegl, R. (2005). SWIFT: A dynamical model of saccade generation during reading. *Psychological Review*, *112*, 777–813. doi:10.1037/0033-295X. 112.4.777

Gil-López, C., Perea, M., Moret-Tatay, C., & Carreiras, M. (2011). Can masked priming effects be obtained with words? *Attention, Perception, and Psychophysics*, *73*, 1643–1649. doi:10.3758/s13414-011-0174-y

Gomez, P., Perea, M., & Ratcliff, R. (2013). A diffusion model account of masked versus unmasked priming: Are they qualitatively different? *Journal of Experimental Psychology: Human Perception and Performance*, *39*, 1731–1740. doi:10.1037/ a0032333

Levy, R., Bicknell, K., Slattery, T., & Rayner, K. (2009). Eye movement evidence that readers maintain and act on uncertainty about past linguistic input. *Proceedings of the National Academy of Sciences*, *106*, 21086–21090. doi:10.1073/pnas. 0907664106

Manso De Zuniga, C., Humphreys, G. W., & Evett, L. J. (1991). Additive and interactive effects of repetition, degradation, and word frequency in the reading of handwriting. In D. Besner & G. W. Humphreys (Eds.), *Basic processes in reading: Visual word recognition* (pp. 10–33). Hillsdale: Erlbaum.

Perea, M., & Acha, J. (2009). Space information is important for reading. *Vision Research*, *49*, 1994–2000. doi:10.1016/j. visres. 2009.05.009

Perea, M., Gil-López, C., Beléndez, V., & Carreiras, M. (2016). Do handwritten words magnify lexical effects in visual word recognition? *The Quarterly Journal of Experimental Psychology*, *69*, 1631–1647. doi:10.1080/17470218.2015. 1091016

Perea, M., & Pollatsek, A. (1998). The effects of neighborhood frequency in reading and lexical decision. *Journal of Experimental Psychology: Human Perception and Performance*, *24*, 767–779. doi:10.1037/0096-1523.24.3.767

Qiao, E., Vinckier, F., Szwed, M., Naccache, L., Valabrègue, R., & Dehaene, S. , et al. (2010). Unconsciously deciphering handwriting: Subliminal invariance for handwritten words in the visual word form area. *NeuroImage*, *49*, 1786–1799. doi:10. 1016/j.neuroimage.2009.09.034

Raney, G. E., & Rayner, K. (1995). Word frequency effects and eye movements during two readings of a text. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale*, *49*, 151–173. doi:10.1037/1196-1961. 49.2.151

Rayner, K., Pollatsek, A., Ashby, J., & Clifton, C. E. (2012). *The psychology of reading*. London: Psychology Press. doi:10.4324/ 9780203155158

Rayner, K., Reichle, E. D., Stroud, M. J., Williams, C. C., & Pollatsek, A. (2006). The effect of word frequency, word predictability, and font difficulty on the eye movements of young and older readers. *Psychology and Aging*, *21*, 448–465. doi:10. 1037/0882-7974.21.3.448

Reichle, E. D. (2015). Computational models of reading: A primer. *Language and Linguistics Compass*, *9*, 271–284. doi:10.1111/lnc3.12144

Reichle, E. D., Pollatsek, A., Fisher, D. L., & Rayner, K. (1998). Toward a model of eye movement control in reading. *Psychological Review*, *105*, 125–157. doi:10.1037/0033-295x. 105.1.125

Reichle, E. D., & Sheridan, H. (2015). E-Z Reader: An overview of the model and two recent applications. In A. Pollatsek & R. Treiman (Eds.), *Oxford handbook of reading* (pp. 277–292). Oxford: Oxford University Press.

Reichle, E. D., Warren, T., & McConnell, K. (2009). Using E-Z Reader to model the effects of higher level language processing on eye movements during reading. *Psychonomic Bulletin and Review*, *16*,1–21. doi:10.3758/pbr.16.1.1

Slattery, T. J. (2009). Word misperception, the neighbor frequency effect, and the role of sentence context: Evidence from eye movements. *Journal of Experimental Psychology:Human Perception and Performance*, *35*, 1969–1975. doi:10. 1037/ a0016894

Slattery, T. J. (2016). Eye movements: From psycholinguistics to font design. In M. C. Dyson & Ch. Y. Yuen (Eds.), *Digital fonts and reading* (pp. 54–78). Singapore: World Scientific. doi:10. 1142/9789814759540_0004

Slattery, T. J., & Rayner, K. (2010). The influence of text legibility on eye movements during reading. *Applied CognitivePsychology*, *24*, 1129–1148. doi:10.1002/acp.1623

White, S. J., Johnson, R. L., Liversedge, S. P., & Rayner, K. (2008). Eye movements when reading transposed text: The importance of word-beginning letters. *Journal of Experimental Psychology: Human Perception and Performance*, *34*, 1261 – 1276. doi:10.1037/0096-1523.34.5.1261