Vision Research 49 (2009) 1994–2000

Contents lists available at ScienceDirect

Vision Research

journal homepage: www.elsevier.com/locate/visres

Space information is important for reading

Manuel Perea *, Joana Acha

Departamento de Metodología, Universitat de València, Facultad de Psicología, Av. Blasco Ibáñez, 21, 46010-Valencia, Spain

ARTICLE INFO

Article history:
Received 3 November 2008
Received in revised form 13 May 2009

Keyword:
Reading

ABSTRACT

Reading a text without spaces in an alphabetic language causes disruption at the levels of word identification and eye movement control. In the present experiment, we examined how word discriminability affects the pattern of eye movements when reading unspaced text in an alphabetic language. More specifically, we designed an experiment in which participants read three types of sentences: normally written sentences, regular unspaced sentences, and alternating**bold** unspaced sentences. Although there was a reading cost in the unspaced sentences relative to the normally written sentences, this cost was much smaller in alternating**bold** unspaced sentences than in regular unspaced sentences.

1. Introduction

Unlike nonalphabetic scripts such as Chinese or Japanese, the vast majority of alphabetic languages employ space information to delimit words. In the past decades, a number of studies have been devoted to the function and importance of interword spaces in normal silent reading (e.g., Bai, Yan, Liversedge, Zang, & Rayner, 2008; Epelboim, Booth, Ashkenazy, Taleghani, & Steinman, 1997; Inhoff, Liu, Wang, & Fu, 1997; Kajii, Nazir, & Osaka, 2001; Kohsom & Gobet, 1997; Malt & Seamon, 1978; Morris, Rayner, & Pollatsek, 1990; Pollatsek & Rayner, 1982; Rayner, Fischer, & Pollatsek, 1998; Sainio, Hyönä, Bingushi, & Bertram, 2007; Spragins, Lefton, & Fisher, 1976). One highly consistent finding is that reading a text without spaces in an alphabetic language causes a reading disruption at the levels of word identification and saccade programming: average fixation durations are longer, the initial landing positions in the words from unspaced sentences is shifted to the beginning of the words, and saccades onto upcoming words (and within a word) are shorter. Two (non-exclusive) explanations have been put forward (see Rayner et al., 1998): (1) reading without spaces makes it difficult to determine where each word begins and ends –an essential component for word identification, and (2) removal of space information makes it difficult to locate the present word and to program the saccade to the same/next word – an essential component for eye movement guidance.

In the present paper, we examine the role of word discriminability when space information is removed in an alphabetic language. By the term "word discriminability", we refer to a quality of the stimulus (i.e., the word) that enables its perceptual discriminability (see Henderson, 2003). We do so by using a new manipulation: each other word was highlighted (using an alternating**bold** manipulation) so that its word boundaries were well-defined (e.g., as in the unspaced sentence The**truth**is**rarely**pure**and**never**simple**). The rationale here is that this manipulation makes it (relatively) easy to determine where the word begins/ends; thereby if word identification is hindered by lack of word "parsing" in regular unspaced sentences, then any effects at the word identification stage (e.g., the word-frequency effect) should be much similar in size in highlighted, unspaced sentences and in normally written sentences. Likewise, the highlighting manipulation should facilitate the location of the present/next word. If so, the pattern of eye movements in alternating**bold** unspaced sentences (e.g., initial landing position, length of within-word saccades) should be close to that of normally written sentences. We should note here that the present manipulation has some resemblance to the one employed by Bai et al. (2008; Experiment 2) in a recent study on Chinese reading. Specifically, Bai and colleagues highlighted the background of each other word (roman vs. italic), as in the sentence The**truth**is**rarely**pure**and**never**simple**.

As indicated above, the role of interword spaces in reading has been examined in a number of studies. Particularly relevant for the present experiment is the seminal study of Pollatsek and Rayner (1982): spaces in text were filled with different types of characters (letters, digits, and blob-like gratings). The more similar the space filler was to a real letter, the more likely it was to disrupt both word identification and eye movement guidance mechanisms. Note that this suggests that readers are – to some degree – sensitive to word discriminability. In a later study designed to separate the effects of spacing on word identification and the effects on eye movement control, Rayner et al. (1998) manipulated the frequency of a target word in spaced and unspaced sentences. (Keep in mind that word-frequency is an important index of the ease or difficulty of word identification, and readers fixate longer on LF words than

* Corresponding author. Fax: +34 963864697.
E-mail address: mperea@valencia.edu (M. Perea).

0042-6989/$ - see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.visres.2009.05.009

on HF words, e.g., see Rayner & Duffy, 1996; White, 2008).[1] In terms of the additive factor logic, if spacing does not affect word identification, then the word-frequency effect should be similar in magnitude for the words embedded in spaced and unspaced sentences (i.e., an additive effect of frequency and spacing). Interestingly, Rayner et al. (1998) found that the removal of space information exaggerated dramatically the word-frequency effect that normally exists with normally spaced text (123 vs. 41 ms), thus strongly suggesting that word identification was hindered by the absence of spaces between words. Furthermore, Rayner et al. (1998) observed that the initial landing on words in unspaced sentences was substantially closer to the beginning of the word than in normally written sentences, and that the size of saccades were shorter for the unspaced sentences – this implies that eye guidance was also hindered.

Obviously, the role of space information in reading can be studied either by removing spaces in alphabetic languages (as in the present experiment), or by adding spaces in languages in unspaced languages like Chinese or Japanese. For instance, Bai et al. (2008) employed a condition in which Chinese sentences were presented with spaces across words (Experiment 1) or in which the sentences were unspaced but the words were highlighted (Experiment 2). The pattern of data in the Bai et al. experiments showed that these two conditions produced an eye movement pattern remarkably similar to that of normal, unspaced Chinese text (see also Sainio et al. 2007, for a similar pattern with Kanji–Kana sentences in Japanese). In addition, there was a reading cost when the spaces were presented within words or when the highlighted text did not correspond to the words boundaries. Similarly, in regular sentences in Japanese – an unspaced language which combines idiographic Kanji with syllabic Kana (e.g., as in the sentence 確かに条件は悪いがつべこべ言うな)–) – there is no advantage of spaced versus unspaced text. (Note that Kanji and Kana characters are visually different: 確 is a ideographic Kanji character, whereas か is the Kana character corresponding to the syllable /ka/.) Importantly, there was an advantage for the spaced text when the sentences only employed syllabic Kana (Sainio et al., 2007; see also Kohsom & Gobet, 1997, for evidence of a similar effect in Thai – another unspaced language). Thus, the differential effect of spacing information in Kanji–Kana vs. only-Kana sentences in Japanese suggests that word discriminability plays a role in eye movement guidance: when visual cues are not available (as in the case of only-Kana sentences), space information provides relevant cues on where to look next. In contrast, space information is not as decisive when words can be "parsed" via visual cues – as in the case of the mixed Kanji–Kana sentences.

In the present experiment, we examined how word discriminability (via an alternating**bold** manipulation) affects normal reading when reading unspaced text in an alphabetic language (Spanish). Specifically, we designed an experiment in which participants read three types of sentences while their eye movements were monitored (see Table 1 for examples): (1) normally written sentences, (2) regular unspaced sentences, and (3) alternating**bold** unspaced sentences. To disentangle the effects on word identification and eye movement guidance, we also manipulated the frequency of a critical target word within the sentence (see Rayner et al., 1998, for similar logic). Thus, for any given sentence, either a high-frequency (HF) or low-frequency (LF) word was embedded in a location of the sentence, such that either word was syntactically and semantically acceptable. This way, it is possible to obtain direct evidence on whether it is lack of space information or, rather, word distinctiveness the key factor that interferes with word identification: if the magnitude of the word-frequency effect is similar for words in normally written sentences and alternating**bold** unspaced sentences, the key factor will be word discriminability. Alternatively, if the word-frequency effect in alternating**bold** sentences is close to that in regular unspaced sentences, the key factor will be lack of space information.

Finally, we also analysed the impact of the alternating**bold** manipulation on eye movement control: if word discriminability makes it easier to program the saccade to the next word, then landing positions in alternating**bold** unspaced sentences should be closer to the preferred viewing location (i.e., between the beginning and the center of the word; Rayner, 1979) than the landing positions in regular unspaced sentences, and the size of within-word saccades in alternating**bold** unspaced sentences should be similar in magnitude to that in normally written sentences.

Table 1
Illustration of the spacing conditions in the experiment.

Spacing condition	Example sentence.
Normal	The yellow car is parked at the corner
Unpaced, regular	Theyellowcarisparkedatthecorner
Unspaced, alternating**bold**	The**yellow**car**is**parked**at**the**corner**

2. Method

2.1. Participants

Twenty-four psychology students from the Universitat de València took part in the experiment and received course credit. All participants had normal vision and were native speakers of Spanish. None of them reported having any speech/reading problems. They were all naïve as to the purpose of the experiment.

2.2. Materials

The stimuli comprised two sets of 60 pairs of sentences in Spanish (see http://www.uv.es/mperea/P&AVisRes.pdf). The members of each sentence pair were identical except for the target word. In sixty of these sentences, a LF noun was inserted in the target location, and in the other sixty sentences, a HF noun was inserted in the target location. Two sentences were made for each target word to minimize a potential effect of context and also to make it possible to test the same words across two different sentences. For example, one sentence frame was "La alumna podrá dominar el violín/idioma con trabajo constante" and the parallel frame was "La profesora sabe manejar el violín/idioma con mucha destreza" ("violín" [the Spanish for violin] is the LF target, whereas "idioma" [the Spanish for language] is the HF target.) The participants who read the first sentence frame with the LF target would then read the second sentence frame with the HF target – or the other way around. LF and HF words were matched in length ($M = 6$ letters, range: 5–9) and the mean word-frequency of LF and HF word targets were 4.5 and 87.3 per million, respectively (range: 0.2–20 and 23–353 for the LF and HF words, respectively; Davis & Perea, 2005). The target words were of low predictability (i.e., they were predicted less than 5% of the time in a "cloze" task conducted prior to the experiment with another sample of subjects) and the sentences had a maximal length of 62 characters so that they fit on a single line of the display. Target words were always around the middle of each sentence (fourth or fifth word position).

[1] Another index of the ease/difficulty of word identification is the number of higher frequency "orthographic" neighbors (e.g., the processing of the word spice is slowed down because of the higher frequency neighbor space; see Acha & Perea, 2008; Davis, Perea, & Acha, in press; Pollatsek, Perea, & Binder, 1999).

2.3. Design

Six lists were created, each containing 120 sentences. Forty sentences were presented in normal text, forty sentences in regular unspaced text, and forty sentences in alternating**bold** unspaced text (see Table 1). The sentences were counterbalanced across the six lists, so that the corresponding LF and HF target words were included in all conditions across the two sets of sentences. The order of the sentences was randomized for each participant.

2.4. Apparatus

The eye movements of the participants were recorded with an EyeLink II eye tracker manufactured by SR Research Ltd. (Canada). The sampling rate for the pupil size and location is of 500 Hz. The average gaze position error is less than 0.5°, and access to eye position data is done only with a 3-ms delay. Registration was binocular, although only data from the right eye was analysed. The position of the participant with respect to the screen was controlled by a head-tracking camera that served for compensating possible head motion.

2.5. Procedure

Each participant was tested individually. Participants completed the experiment in a well-lit soundproof room. Participants were sitting in a chair that ensured a distance of 75 cm from the center of the screen. After the calibration and validation process, participants read eight practice sentences. Each trial started with the presentation of a fixation point that was left aligned (coinciding with the location of the first letter of each sentence). When the fixation point disappeared from the screen, the target sentence was displayed. Participants were instructed to read for comprehension and to press one button on a gamepad as soon as they finished reading the sentence. To assure comprehension, they were asked to answer comprehension questions about the sentence they had just read after 20% of the sentences. Participants had little difficulty answering the questions correctly (over 94% of correct responses; there were no significant differences across experimental conditions).

2.6. Data analysis

Across all the trials, approximately 3% of the data were lost due to a track loss. The remaining data were analysed first with respect to global performance characteristics such as reading rate, total reading time and fixation duration and then local analyses of eye behavior on the target words were conducted to test specific hypotheses. Fixations under 80 ms that were within one letter of the next or previous fixation were merged into that fixation. Any remaining fixations below 80 ms or over 1200 ms were discarded. Repeated measures Analyses of Variance (ANOVAs) based on participant variability were undertaken. List was included in all the statistical analyses to extract the variability due to the counterbalancing lists (Pollatsek & Well, 1995).[2]

[2] We only report F ratios over participants (i.e., F1 ratios). This is the appropriate analysis for testing the significance of the effects in a counterbalanced design, such as that used in the present study (see Raaijmakers, 2003). In any case, the p values corresponding to the F2 ratios (i.e., "item" analyses) in the present experiment essentially mimicked the reported p values of the F1 ratios.

Table 2
Global measures for each of the conditions (mean and standard deviation): total sentence reading time (in ms), average fixation duration (in ms), and number of words per second (reading rate).

	Total reading time		Average fixation duration		Words per second	
	M	*SD*	*M*	*SD*	*M*	*SD*
Spacing condition						
Normal	1856	128	210	4.2	6.1	0.5
Unspaced, alternating bold	2708	126	229	4.3	3.9	0.3
Unspaced, regular	3343	101	244	5.1	3.1	0.2

3. Results

3.1. Global analysis

This analysis provides important temporal information on the differential reading cost across the three spacing conditions. The average reading time, average fixation duration, and reading rate are summarized in Table 2. All the global measures showed a significant effect of Type of script (all *ps* < .001, see Table 3 for *F* values): normally written sentences were read faster and with less overall fixation durations than unspaced sentences. More important, alternating**bold** unspaced sentences were read faster and with fewer fixations than regular unspaced sentences (see Table 3, for the pairwise comparisons).

Thus, these data demonstrate that temporal measures are affected by the spacing manipulation. More important, they also show that visual discriminability of words plays an important role during reading.

3.2. Local measures

Local measures allow us to examine the role of word-frequency of the target words that were embedded in identical sentence frames (see Tables 4 and 5 for reading measures and *F* values, respectively). Furthermore, local analyses provide essential information on whether the spatial characteristics of the pattern of eye movements are affected by the removal of spacing information and/or by the visual cues that mark the boundaries between words (see Tables 6 and 7 for the size of saccade length and *F* values, respectively).

3.2.1. Percentage of skipping the target word

This is the percentage of times in which the readers skipped the target word in the first-pass eye measures. The main effects of frequency and spacing were significant, whereas the interaction between the two factors did not approach significance (see Tables 4 and 5). As usual, target words of high-frequency were skipped more than target words of high-frequency (see O'Regan, 1979; Rayner et al., 1998). In addition, target words were more frequently skipped in normally written sentences than in alternating**bold** unspaced sentences, $F(1, 18) = 32.35$, $p < .001$, and target words were more skipped in the alternating**bold** unspaced condition than in regular unspaced sentences, although this difference only approached statistical significance, $F(1, 18) = 3.49$, $p = .07$. There were no signs of an interaction between the two factors.

3.2.2. First fixation duration

This is the amount of time a reader spends on the initial fixation of the target word. The main effects of frequency and spacing were significant, whereas the interaction between the two factors was not significant (see Tables 4 and 5). First fixation durations were 13 ms shorter for high-frequency words than for low-frequency words. As for the spacing conditions, first fixation durations were

Table 3
Global measures: ANOVA main effects of spacing and pairwise comparisons. *P* values are less than the stated values.

	Total reading time			Average fixation duration			Words per second		
	F	*MSE*	*P*	*F*	*MSE*	*P*	*F*	*MSE*	*P*
ANOVA									
Spacing	314.14	44223.2	.001	112.2	60.4	.001	124.17	0.4	.001
Pairwise comparisons									
Normal vs. Unspaced (Regular)	382.64	78054.8	.001	122.72	110.1	.001	129.55	0.89	.001
Normal vs. Unspaced (Alt.Bold)	417.47	21924.5	.001	83.24	51.5	.001	118.67	0.5	.001
Unsp. (Alt.Bold) vs. Unsp. (Regular)	128.00	38690.2	.001	129.41	19.8	.001	100.94	0.1	.001

Note: The degrees of freedom were (2, 36) for the ANOVA and (1, 18) for the pairwise comparisons.

18 ms shorter in normally written sentences that in alternating**bold** unspaced sentences, $F(1, 18) = 21.38$, $p < .001$, and first fixation durations were 16 ms sorter in alternating**bold** unspaced sentences than in regular unspaced sentences, $F(1, 18) = 21.28$, $p < .001$.

3.2.3. Gaze duration on the target word

Gaze duration represents the sum of fixation durations on a target word before the reader leaves that word. Again, the main effect of spacing reflected clear differences between the three conditions: normally written sentences, then alternating**bold** unspaced sentences, and finally, regular unspaced sentences (see Tables 4 and 5). There was also an effect of word-frequency. More important, the interaction between spacing and frequency was robust: the word-frequency effect was dramatically greater in regular unspaced sentences (125 ms), $F(1, 18) = 80.60$, $p < .001$, than in the normally written sentences (53 ms), $F(1, 18) = 25.20$, $p < .001$, or in the alternating**bold** unspaced sentences (46 ms), $F(1, 18) = 10.66$, $p < .005$. That is, the effect of word-frequency increases dramatically when information about where the word begins/ends is not available (see Rayner et al., 1998, for similar evidence with regular unspaced sentences).

3.2.4. Total time

The total time spent on the target word corresponds to the sum of all the fixation durations on the target word, including first-pass and regressive fixations. There was a main effect of both spacing and frequency (see Tables 4 and 5 for reading times and *F* values, respectively). As occurred with the gaze durations, there was a significant interaction between spacing and frequency: Low-frequency words were read 165 ms slower than high-frequency words in regular unspaced sentences, $F(1, 18) = 52.06$, $p < .001$, whereas the word-frequency effect was reduced to 105 ms in alternating**bold** unspaced sentences, $F(1, 18) = 38.51$, $p < .001$, and to 66 ms in normally written sentences, $F(1, 18) = 24.89$, $p < .001$. (Note that, unlike gaze durations, the size of the word-frequency effect was larger for words embedded in alternating**bold** unspaced sentences than for words embedded in normally written sentences, as shown by the significant interaction when only these two spacing conditions are included in the ANOVA, $F(1, 18) = 6.19$, $p < .025$).

3.2.5. Percentage of regression

Here we examine the percentage of regressions back to the target word (see Tables 4 and 5). We found that the number of regressions back to the target word was significantly greater for low-frequency words than for high-frequency words, and we also found a main effect of spacing: the percentage of regressions back to the target word in the alternating**bold** condition was significantly greater than in the normal condition, $F(1, 18) = 41.19$, $p < .007$, whereas we found no signs of a difference between the percentage of regressions in the two unspaced conditions ($F < 1$). This suggests that reading without spaces slows down a verification/late stage of lexical access. There were no trends of an interaction between the two factors.

3.2.6. Initial landing position

The initial landing position on the target word of each sentence was assessed for all target words. To make observations from words of different length comparable, the analysis was based on the subdivision of the words into five fixation zones (see Rayner & Fischer, 1996; Rayner et al., 1998; Vitu, O'Regan, Inhoff, & Topolski, 1995, for a similar procedure). In normally written sentences, the landing location followed the usual pattern of the preferred viewing position (Rayner, 1979), that is, the initial landing position was a bit to the left of the center of the word. The location of the first fixation of the target word in alternating**bold** unspaced sentences was more to the beginning of the word than in normally written sentences (2.1 vs. 2.3, respectively, $F(1, 18) = 15.34$, $p < .001$) and, in turn, the location of the first fixation in regular unspaced sentences was more to the beginning of the word than in alternating**bold** unspaced sentences (1.8 vs. 2.1, respectively, $F(1, 18) = 11.70$, $p < .003$). This occurred regardless the frequency of the word – we found no signs of a main effect of frequency or an interaction between the two factors (see Tables 4 and 5).

3.2.7. Effect of initial landing position on subsequent processing of the target word

As noted by Rayner et al. (1998), some of the observed effects on the local analyses (first fixation durations and gaze durations, in particular) could have been influenced by the differential pattern of landing position in unspaced sentences. To control for landing

Table 4
Local measures for the different experimental conditions in the experiment (mean and standard deviation): first fixation duration, gaze duration, total time, location of the first fixation, % of regressions, and % of word skipping.

	First fix. duration Frequency				Gaze duration Frequency				Total time Frequency				Fixation location Frequency				% Regression Frequency				% Skipping Frequency			
	Low		High		Low		High		Low		High		Low		High		Low		High		Low		High	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Spacing																								
Normal	223	16	204	14	278	46	225	13	338	57	272	35	2.4	0.3	2.3	0.2	14.4	1.4	12.2	1.5	14.8	1.0	17.2	1.3
Unspaced, alt.bold	234	23	230	21	368	55	322	41	494	78	389	48	2.1	0.3	2.1	0.2	21.0	1.0	16.4	1.2	5.0	1.0	9.8	1.3
Unspaced, regular	255	25	242	13	503	81	378	49	671	104	506	69	1.8	0.2	1.9	0.2	20.5	1.2	19.2	1.1	4.1	1.1	6.6	1.3

Table 5
Local measures: 3 (spacing: normal, regular unspaced, alternating bold unspaced) × 2 (frequency: high, low) ANOVA. *P* values are less than the stated values unless preceded by=.

Local measures	Total time			First fix. duration			Gaze duration			Fixation location			% Regression			% Skipping		
Effect	*F*	*MSE*	*p*	*F*	*MSE*	*p*	*F*	*MSE*	*p*	*F*	*MSE*	*p*	*F*	*MSE*	*p*	*F*	*MSE*	*p*
Spacing	208.85	4614.3	.001	37.95	371.9	.001	122.37	3502.5	.001	23.56	.1	.001	7.80	77.3	.002	25.78	64.9	.001
Frequency	87.53	5139.8	.001	15.34	269.3	.001	66.07	3028.3	.001	0.74	.05	= .39	5.15	76.3	.03	9.06	40.5	.007
Spacing × Frequency	9.11	3315	.001	2.79	243.4	=.06	15.40	1500.0	.001	1.89	.08	=.16	1.59	45.8	=.3	0.30	29.5	=.70

Note: The degrees of freedom were (1, 18) for the main effect of Frequency and (2, 36) for the other two effects.

position, the target words were divided into five critical zones (see above), and we included Initial landing position as a factor in the ANOVA – together with spacing and frequency. Because there were several empty cells in the two rightmost zones for some participants, our analyses only included the first three zones.

For first fixation durations, the pattern of data mimicked that of the local analyses. That is, we found a main effect of spacing, $F(2, 36) = 26.30$, $p < .001$, and a main effect of frequency, $F(1, 18) = 11.58$, $p < .001$, whereas the interaction between these two factors was not significant, $F(2, 36) = 2.09$, $p > .13$. As in the Rayner et al. (1998) experiments, there were no trends of an effect of Initial Landing position, $F < 1$, or of any interactions of Initial Landing position with spacing or frequency.

For gaze durations, the pattern of data also mimicked that of the local analyses. That is, we found a main effect of spacing, $F(2, 36) = 81.06$, $p < .001$, a main effect of frequency, $F(1, 18) = 35.30$, $p < .001$, and a significant interaction between spacing and frequency, $F(2, 36) = 6.61$, $p < .005$. Again, this interaction reflected an exaggerated frequency effect for the regular unspaced sentences (117 ms) as compared to the alternating**bold** unspaced sentences and normally written sentences (34 and 56 ms, respectively). As in the Rayner et al. (1998) experiments, we found a robust effect of Initial Landing position, $F(2, 36) = 23.56$, $p < .001$ (average fixation durations of 384, 339, and 316 ms – across the three spacing conditions – for the first, second, and third zone, respectively), and no signs of any interactions of Initial Landing position with spacing or frequency (all *ps* > .30).

In sum, the pattern of data reported in the local analyses was not an artifact of the initial landing position. Unsurprisingly, an analysis of the pattern of refixations mimicked that of gaze durations – this explains why the effect of initial landing position occurred for gaze durations but not for first fixation durations (see Rayner et al., 1998, for a similar pattern). We should note here that we also examined the N-1 fixation durations; as in the Rayner et al. (1998) experiment, we failed to find any signs of a word- frequency effect on the N-1 fixations. Finally, we examined the duration of the first fixation after the reader leaves the target word (*spillover* effect) and there were no signs of a word-frequency effect (4 ms for normally written sentences, 0 ms for alternating**bold** sentences and 1 ms for regular unspaced sentences); bear in mind that spillover effects are often a noisy measure because the first fixation after the reader leaves the target word can be either on the word following the target or on the word following that (see Acha & Perea, 2008; White & Liversedge, 2006, for a failure to obtain spillover effects in lexical factors).

3.2.8. Size of saccades in target region

We also examined the length of saccades entering, leaving, and within the target word (see Tables 6 and 7 for reading times and *F* values, respectively). Here the saccade length is given in absolute units (character spaces) because attempting to adjust for the length of the target word would be more difficult to interpret (Rayner et al., 1998).

The length of the saccades within words was shorter in regular unspaced sentences (3.6 characters long) than in normally written sentences and alternating**bold** unspaced sentences (4.0 characters in the two cases), $F(1, 18) = 7.65$, $p < .01$, and $F(1, 18) = 19.60$, $p < .001$, respectively. In addition, we found a longer (forward) saccade length for high-frequency words than for low-frequency words (see Tables 6 and 7), which suggests that lexical information (i.e., word-frequency) is available very early during processing – before the decision of where to move the eyes (see Reichle, Pollatsek, Fisher, & Rayner, 1998). With regard to the size of backward saccades within words, the length of these saccades was shorter for regular unspaced sentences (2.8 characters) than for alternating**bold** unspaced sentences (3 characters) although the difference did not reach the criterion for statistical significance ($F(1, 18) = 3.95$, $p = .06$), and in turn, the length of the saccades for normally written sentences was not significantly longer than for unspaced alternating**bold** sentences ($F(1, 18) = 2.39$, $p > .13$). (The length of backward saccades within words was longer for normally written sentences than for regular unspaced sentences, $F(1, 18) = 9.03$, $p < .009$.) Finally, there were no signs of an interaction of spacing and frequency in any of the within-word saccade length measures.

With respect to the forward saccades entering and leaving the target word, the length of the saccades into the word was longer in normally written sentences (8.3 characters) than in alternating**bold** unspaced sentences (6.0 characters), $F(1, 18) = 132.71$, $p < .001$, and, in turn, the length of the saccades was longer in alternating**bold** unspaced sentences than in regular unspaced sentences (6.0 vs. 5.0 characters, respectively), $F(1, 18) = 26.11$, $p < .001$. Likewise, the length of the saccades out of the target word was significantly longer in normally written sentences (10.4 characters) than in alternating**bold** unspaced sentences (7.5 characters), $F(1, 18) = 336.58$, $p < .001$; and, in turn, the length of the saccades out of the target word was significantly longer in alternat-

Table 6
Local saccade measures for each of the experimental conditions (in characters): mean within word forward saccade length, mean within word regressive saccade length, mean length of forward saccade into target word, mean length of forward saccade out of target word.

	Forward within saccade length Frequency				Regressive within saccade length Frequency				Length of saccade into target word Frequency				Length of saccade out of target word Frequency			
	Low		High		Low		High		Low		High		Low		High	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Spacing																
Normal	3.9	0.5	4.2	0.6	2.5	0.2	3.9	0.5	8.4	0.6	8.2	0.7	10.2	1.0	10.6	0.7
Unspaced, alt.bold	3.8	0.4	4.2	0.4	2.7	0.5	3.2	0.7	5.8	0.7	6.1	0.6	7.5	0.4	7.5	0.5
Unspaced, regular	3.5	0.3	3.7	0.3	2.9	0.4	2.6	0.3	4.8	0.5	5.2	0.5	6.2	0.6	6.7	0.7

Table 7
Local saccade measures: 3 (Spacing: normal, regular unspaced, alternating bold unspaced) × 2 (frequency: high, low) ANOVA. Main effect and interactions. *P* values are less than the stated values unless preceded by =.

	Forward saccade length			Regressive saccade length			Length of saccade into target word			Length of saccade out of target word		
Effect	*F*	*MSE*	*p*	*F*	*MSE*	*p*	*F*	*MSE*	*p*	*F*	*MSE*	*p*
Spacing	5.96	50.8	.006	5.66	169.7	.007	113.28	124.7	.001	204.4	100.8	.001
Frequency	14.41	19.3	.001	6.70	280.6	.018	0.89	39.1	=.35	18.3	17.1	.001
Spacing × Frequency	0.28	16.3	=.17	1.47	179.8	=.26	2.06	41.1	=.14	1.80	45.7	=.17

Note: The degrees of freedom were (1, 18) for the main effect of Frequency and (2, 36) for the other two effects.

ing**bold** unspaced sentences than in regular unspaced sentences (7.5 vs. 6.5 characters, respectively), $F(1, 18) = 29.23$, $p < .001$. Neither the main effect of frequency nor the interaction between frequency and spacing condition was significant in any of the two across-words saccade length measures.

In sum, the size of the saccades within words for alternating**bold** unspaced sentences and normal sentences was fairly similar, whereas the size of the saccades for regular unspaced sentences was shorter. With respect to the across-words saccades, the size of these saccades is longer in normally spaced than in unspaced sentences. Of course, we have to take into account that words in unspaced sentences are closer to each other than in spaced sentences, and this makes it difficult to make strong claims when comparing spaced and unspaced sentences. In any case, the length of the saccades was longer for alternating**bold** unspaced sentences than for regular unspaced sentences.

4. Discussion

The present experiment examined the role of word discriminability on the eye movement pattern when reading unspaced text in an alphabetic language. Not surprisingly, when spaces were removed completely, the reading rate decreased dramatically and the eye movement pattern differed from that in normal reading, replicating earlier research (e.g., Rayner et al., 1998). More important, when spaces were removed but visual cues about word boundaries were provided, there was still some reading cost, but the process of word identification was relatively unhindered. We examine the implications of these findings in the following paragraphs.

Importantly, when there is a physical cue for where a word begins/ends in unspaced sentences, the stage of word identification is not severely hampered. Indeed, if one looks at the gaze duration data, the magnitude of the word-frequency effect was remarkably similar for normally written sentences and alternating**bold** unspaced sentences, whereas it was dramatically higher in regular unspaced sentences – thus replicating and extending the findings reported by Rayner et al. (1998). That is, gaze durations were lengthened by low-frequency words in the regular unspaced sentences – but not in the alternating**bold** unspaced sentences. This indicates that the most relevant factor is not lack of space information per se, but rather it is the lack of information on the beginning/end of the word. This is consistent with the recent experiment of Sainio et al. (2007) in Japanese – an unspaced language. Sainio et al. found that "syllabic" Kana sentences benefited from space information, whether the regular Kanji–Kana sentences did not (presumably because in the latter case, the reader has better visual cues to "parse" the beginning/end of the words).

With respect to eye guidance, word discriminability also plays a role in eye movement programming. The pattern of within-word saccades observed in alternating**bold** unspaced sentences is very close to that of the normally written sentences, and the initial landing position in alternating**bold** unspaced sentences is close to that of normally written sentences – at least much closer than that of regular unspaced sentences. Of course, the size of across-word saccades is substantially longer for the normally written sentences than for the alternating**bold** unspaced sentences, but in the latter case words are visually closer (i.e., readers do not have to skip any spaces). Thus, the information provided by the alternating**bold** unspaced sentences goes beyond a mere visual cue for making word identification easier: it also helps eye movement programming. However, the pattern of eye movements in unspaced alternating**bold** sentences also showed relevant differences relative to normally written sentences, as reflected in the percentage of skipping words and the percentage of regressions.

One might argue that the pattern of eye movements in the alternating**bold** condition due to the greater visual saliency (rather than visual discriminability) of the **bold** words. As an anonymous reviewer indicated, the bold words in the alternating**bold** condition may command more attention than the non-bold text. If so, the obtained pattern of data effect would reflect visual saliency rather than visual discriminability of the bold text. To examine this potential explanation, we conducted post hoc pairwise comparisons for the "bold" and the "normal" target words in the alternating**bold** condition. (In the experiment, the target word in the alternating**bold** sentences was bold in one list, and normal in the other.) Results showed that the critical values were remarkably similar when the targets were written in bold or not (all ps > .23). Thus, the difference between the normal and alternating**bold** conditions was not due to the demanding attention of the perceptual saliency of the bold text.

Therefore, two relevant conclusions can be derived from this experiment: (1) word discriminability (via an alternating**bold** manipulation) aids the process of visual-word identification – as deduced from the word-frequency data, (2) space information aids in locating words in text and it is important for eye guidance control across words – as deduced from the eye movement pattern.[3] What should be noted here is that the present data nicely replicate and extend the findings reported by Rayner et al. (1998). How can theoretical models explain the present pattern of data? First, we will analyse the present findings in terms of a highly influential model of eye movement control, namely the E–Z Reader model (Reichle et al., 1998; see also Pollatsek, Reichle, & Rayner, 2006; Rayner, Ashby, Pollatsek, & Reichle, 2004). In the E–Z Reader model, decisions about where to fixate next are determined largely by low-level visual cues in the text, such as the interword spaces (i.e., low-spatial frequency information), whereas decisions of when to move the eyes are influenced (mostly but non-exclusively) by the ease/difficulty associated with processing a word (via high-spatial information that is processed by the word identification module). It is reasonable to assume that in the alternating**bold** sentences, the low-level information from the differing format across words helped (to some degree) the decisions of where to fixate next – as reflected by the

[3] We acknowledge that as with any "text" manipulation, the present experiment is not free from potential shortcomings. One potential reason for the disadvantage of the alternating**bold** unspaced sentences relative to normally written sentences is lateral interference: external letters in unspaced sentences may be laterally masked by their contiguous letters (e.g., see Bouma, 1973).

longer saccades relative to the regular unspaced sentences. Furthermore, as deduced from the similar size of the word-frequency effect in the target words embedded in normal sentences and alternating**bold** unspaced sentences, the initial index of the word's familiarity (stage L_1 in the model), which serves as the trigger for programming an eye movement towards the next word was not severely hindered. In addition, the fact that the percentage of regressions back to the target word was similar in the two unspaced conditions –and substantially greater than in the normally written condition – strongly suggests that the completion of word identification (stage L_2) was more seriously hindered by lack of interword spaces than the L_1 stage. Finally, the present data showed no signs of a frequency effect on N-1 fixations, as predicted by the E–Z Reader model. It should be stressed that the lack of a frequency effect on N-1 fixations is identical to that reported by Rayner et al. and consistent with the lack of a parafoveal-on-foveal effect (see also Rayner & Juhasz, 2004; Starr & Rayner, 2003, for similar findings). Although the present set of data is complex, we believe that the E–Z Reader data may provide a nice approximation to how the eye movement patterns vary across spaced vs. unspaced text (see Rayner, Reichle, Stroud, Williams, & Pollatsek, 2006, and Rayner, Li, & Pollatsek, 2007, for successful application of the E–Z Reader model to an unfamiliar font format and to unspaced Chinese, respectively). Another successful model of eye movement control is the SWIFT model (Engbert, Nuthmann, Richter, & Kliegl, 2005). Although the E–Z Reader model and the SWIFT model differ in a number of core assumptions (e.g., see Reichle, Rayner, & Pollatsek, 2003), the two models share a number of predictions. Thus, it seems plausible that the SWIFT model could also accommodate the observed effects of spaced vs. unspaced text – to our knowledge, no specific simulation work has been performed on how the SWIFT model can simulate unspaced text. (Note, however, that the absence of a parafoveal-on-foveal effect in the present experiment is more consistent with the E–Z Reader model than with the SWIFT model).

To sum up, the present experiment demonstrates that when word discriminability is enhanced in unspaced text – as in alternating**bold** sentences, eye movement measures differ clearly from those in regular unspaced text: words are skipped more often, the amount of time spent on the words is reduced, word-frequency effects have the "usual" magnitude, and eye movements follow a more word-based pattern (e.g., initial landing positions). Thus, one basic responsible for the difficulty in reading without spaces in alphabetic languages relies on the difficulty in "parsing" the words and where to look next (see Rayner, 1997; Rayner et al., 1998). Finally, we would like to note that the present manipulation can be employed not only to delimit words, but also to delimit syllables, lexemes, and/or morphemes in morphologically complex words.

Acknowledgements

The research reported in this article has been partially supported by Grants PSI2008-04069/PSIC and CONSOLIDER-INGENIO2010_CSD2008-00048 from the Spanish Ministry of Science and Innovation. Joana Acha was the recipient of a post-graduate grant (BFI05.33) from the Basque Government. We would like to thank Keith Rayner and an anonymous reviewer for helpful comments on an earlier version of the paper.

References

Acha, J., & Perea, M. (2008). The effect of neighborhood frequency in reading: Evidence with transposed-letter neighbors. *Cognition, 108*, 290–300.

Bai, X., Yan, G., Liversedge, S. P., Zang, C., & Rayner, K. (2008). Reading spaced and unspaced Chinese text: Evidence from eye movements. *Journal of Experimental Psychology: Human Perception and Performance, 34*, 1277–1287.

Bouma, H. (1973). Visual interference in the parafoveal recognition of initial and final letters of words. *Vision Research, 13*, 762–782.

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

Davis, C.J., Perea, M., & Acha, J. (in press). Re(de)fining the orthographic neighborhood: The role of addition and deletion neighbors in lexical decision and reading. *Journal of Experimental Psychology: Human Perception and Performance*.

Engbert, R., Nuthmann, A., Richter, E., & Kliegl, R. (2005). SWIFT: A dynamic model of saccade generation during reading. *Psychological Review, 112*, 777–813.

Epelboim, J., Booth, J. R., Ashkenazy, R., Taleghani, A., & Steinman, R. M. (1997). Fillers and spaces in text: Evidence for the importance of word recognition during reading. *Vision Research, 37*, 2899–2914.

Henderson, J. M. (2003). Human gaze control during real-word scene perception. *Trends in Cognitive Sciences, 7*, 498–504.

Inhoff, A. W., Liu, W., Wang, J., & Fu, D. J. (1997). Use of spatial information during the reading of Chinese text. In D. L. Peng, H. Shu, & H. C. Chen (Eds.), *Cognitive research on Chinese language* (pp. 296–329). Jinan, China: Shan Dong Educational Publishing.

Kajii, N., Nazir, T. A., & Osaka, N. (2001). Eye movement control in reading unspaced text: The case of Japanese script. *Vision Research, 41*, 2503–2510.

Kohsom, C., & Gobet, F. (1997). Adding spaces to Thai and English: Effects on reading. *Proceedings of the Cognitive Science Society, 19*, 388–393.

Malt, B. C., & Seamon, J. G. (1978). Peripheral and cognitive components of eye guidance in filled-space reading. *Perception & Psychophysics, 23*, 399–402.

Morris, R. K., Rayner, K., & Pollatsek, A. (1990). Eye movement guidance in reading. *Journal of Experimental Psychology: Human Perception and Performance, 16*, 268–281.

O'Regan, J. K. (1979). Eye guidance in reading: Evidence for linguistic control hypothesis. *Perception & Psychophysics, 25*, 501–509.

Pollatsek, A., & Rayner, K. (1982). Eye movement control in reading: The role of word boundaries. *Journal of Experimental Psychology: Human Perception and Performance, 8*, 817–833.

Pollatsek, A., & Well, A. (1995). On the use of counterbalanced designs in cognitive research: A suggestion for a better and more powerful analysis. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*, 785–794.

Pollatsek, A., Perea, M., & Binder, K. (1999). The effects of neighborhood size in reading and lexical decision. *Journal of Experimental Psychology: Human Perception and Performance, 25*, 1142–1158.

Pollatsek, A., Reichle, E. D., & Rayner, K. (2006). Tests of the E–Z Reader model: Exploring the interface between cognition and eye-movement control. *Cognitive Psychology, 52*, 1–56.

Raaijmakers, J. G. W. (2003). A further look at the "language-as-fixed-effect fallacy". *Canadian Journal of Experimental Psychology, 57*, 141–151.

Rayner, K. (1979). Eye guidance in reading: fixation locations within words. *Perception, 8*, 21–30.

Rayner, K. (1997). Understanding eye movements in reading. *Scientific Studies of Reading, 1*, 317–339.

Rayner, K., Ashby, J., Pollatsek, A., & Reichle, E. (2004). The effects of frequency and predictability on eye fixations in reading: Implications for the E–Z Reader model. *Journal of Experimental Psychology: Human Perception and Performance, 30*, 720–732.

Rayner, K., & Duffy, S. A. (1986). Lexical complexity and fixation times in reading: Effects of word frequency, verb complexity, and lexical ambiguity. *Memory & Cognition, 14*, 191–201.

Rayner, K., Fischer, M. H., & Pollatsek, A. (1998). Unspaced text interferes with both word identification and eye movement control. *Vision Research, 38*, 1129–1144.

Rayner, K., Li, X., & Pollatsek, A. (2007). Extending the E–Z Reader model of eye movement control to Chinese readers. *Cognitive Science, 31*, 1021–1034.

Rayner, K., Reichle, E. D., Stroud, M. J., Williams, C. C., & Pollatsek, A. (2006). The effects of word frequency, word predictability, and font difficulty on the eye movements of young and elderly readers. *Psychology and Aging, 21*, 448–465.

Rayner, K., & Fischer, M. H. (1996). Mindless reading revisited: Eye movements during reading and scanning are different. *Perception & Psychophysics, 58*, 734–747.

Rayner, K., & Juhasz, B. J. (2004). Eye movements in reading: Old questions and new directions. *European Journal of Cognitive Psychology, 16*, 340–352.

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.

Reichle, E. D., Rayner, K., & Pollatsek, A. (2003). The E–Z Reader model of eye movement control in reading: Comparisons to other models. *Behavioral and Brain Sciences, 26*, 445–476.

Sainio, M., Hyönä, J., Bingushi, K., & Bertram, B. (2007). The role of interword spacing in reading Japanese: An eye movement study. *Vision Research, 20*, 2575–2584.

Spragins, A. B., Lefton, L. A., & Fisher, D. F. (1976). Eye movements while reading and searching spatially transformed text: A developmental examination. *Memory & Cognition, 4*, 36–42.

Starr, M. S., & Rayner, K. (2003). Eye movements during reading: Some current controversies. *Trends in Cognitive Sciences, 5*, 156–163.

Vitu, F., O'Regan, J. K., Inhoff, A. W., & Topolski, R. (1995). Mindless reading: Eye movement characteristics are similar in scanning letter strings and reading texts. *Perception & Psychophysics, 57*, 352–364.

White, S. J. (2008). Eye movement control during reading: Effects of word frequency and orthographic familiarity. *Journal of Experimental Psychology: Human Perception and Performance, 34*, 205–223.

White, S. J., & Liversedge, S. P. (2006). Foveal processing difficulty does not modulate non-foveal orthographic influences on fixation positions. *Vision Research, 46*, 426–437.