Journal of Vision (2011) 11(6):8, 1–21 http://www.journalofvision.org/content/11/6/8

The impact of letter spacing on reading: A test of the bigram coding hypothesis

Fabien Vinckier
ICM Research Center, UMRS 975, INSERM, Paris, France, &
Faculté de Médecine Pitié-Salpêtrière,
Université Pierre-et-Marie-Curie,
IFR 70, Paris, France

Emilie Qiao
Faculté de Médecine Pitié-Salpêtrière,
Université Pierre-et-Marie-Curie,
IFR 70, Paris, France

Christophe Pallier
Cognitive Neuroimaging Unit, INSERM, Gif sur Yvette, France,
CEA, DSV, I2BM, NeuroSpin Center,
Gif sur Yvette, France, &
Collège de France, Paris, France

Stanislas Dehaene
Cognitive Neuroimaging Unit, INSERM, Gif sur Yvette, France,
CEA, DSV, I2BM, NeuroSpin Center,
Gif sur Yvette, France, &
Collège de France, Paris, France

Laurent Cohen
ICM Research Center, UMRS 975, INSERM, Paris, France,
Faculté de Médecine Pitié-Salpêtrière,
Université Pierre-et-Marie-Curie, IFR 70, Paris, France, &
AP-HP, Department of Neurology, Hôpital de la Salpêtrière,
Paris, France

Identifying letters and their relative positions is the basis of reading in literate adults. The Local Combinations Detector model hypothesizes that this ability results from the general organization of the visual system, whereby object encoding proceeds through a hierarchy of neural detectors that, in the case of reading, would be tuned to letters, bigrams, or other letter combinations. Given the increase of receptive fields by a factor of 2 to 3 from one neural level to the next, detectors should integrate information only for letters separated by at most 2 other characters. We test this prediction by measuring the impact of letter spacing on reading, purifying this effect from confounding variables. We establish that performance deteriorates non-linearly whenever letters are separated by at least 2 blank spaces, with the concomitant emergence of a word length effect. We then show that this cannot be reduced to an effect of physical size nor of visual eccentricity. Finally, we demonstrate that the threshold of about 2 spaces is constant across variations in font size. Those results support the hypothesis that the fast recognition of combinations of nearby letters plays a central role in the coding of words, such that interfering with this representation prevents the parallel analysis of letter strings.

Keywords: letter spacing, reading, bigram

Citation: Vinckier, F., Qiao, E., Pallier, C., Dehaene, S., & Cohen, L. (2011). The impact of letter spacing on reading: A test of the bigram coding hypothesis. *Journal of Vision, 11*(6):8, 1–21, http://www.journalofvision.org/content/11/6/8, doi:10.1167/11.6.8.

Introduction

When children start learning to read, they scan letters one at a time, resulting in a strong positive correlation of reading latencies with word length (Aghababian & Nazir, 2000). Over years of training, the ability develops to identify all the letters in a word in parallel, so that the length effect decreases and eventually vanishes in expert readers (Weekes, 1997). According to the local combination detector (LCD) model, such parallel letter encoding is allowed by the fact that, in expert readers, whole words are processed as single visual objects, based on the recycling of neural mechanisms that underlie the perception of complex objects in general (Dehaene, Cohen, Sigman, & Vinckier, 2005). Object encoding takes place

doi: 10.1167/11.6.8 Received January 25, 2010; published May 12, 2011 ISSN 1534-7362 © ARVO

in the ventral visual pathway, through a hierarchy of converging neural detectors with increasingly wider receptive fields, tuned to increasingly larger object parts (Serre, Oliva, & Poggio, 2007). In the case of words, such hierarchically embedded object parts may consist in letter fragments, full letters, bigrams (i.e., pairs of letters), and even larger chunks such as morphemes, for which detectors may develop through intensive training. A subset of the ventral pathway critical to word reading is thought to be located in the left fusiform region, as shown by converging activation studies and lesion data (Cohen et al., 2000; Gaillard et al., 2006).

Whenever parallel letter processing is impeded, due to stimulus degradation (or to left fusiform lesions), adult readers revert to a piecemeal serial reading mode, as revealed by the resurgence of a word length effect (Ellis, 2004). Various types of degradation may yield serial reading: (1) low-level degradation such as low-contrast displays (Legge, Ahn, Klitz, & Luebker, 1997); (2) the use of unfamiliar formats to which the visual system has not been trained, e.g., mIxEd case (Lavidor, 2002), vertically printed words (Bub & Lewine, 1988), or words displayed in the left visual field (Lavidor & Ellis, 2002); (3) the insertion of blank space between consecutive letters (Cohen, Dehaene, Vinckier, Jobert, & Montavont, 2008), which is the focus of the present study. Whatever the degradation method, readers must serially attend to letters or word fragments whenever letters cannot be effectively processed in parallel over the whole string (Cohen & Dehaene, 2009).

Degradation by means of letter spacing deserves to be singled out, as it gives a window into a core feature of visual word perception. Indeed, a consequence of the hierarchical organization of the visual cortex is that object parts can be chunked together into a single perceptual object only if they are sufficiently close together. A principle of retinotopic organization permeates throughout the visual system (Hasson, Levy, Behrmann, Hendler, & Malach, 2002): neurons respond to stimuli within a local "receptive field," support or compete with other neurons coding for nearby locations through medium-range horizontal connections, and project to hierarchically higher areas in a retinopy-preserving manner, such that receptive fields broaden by a factor of 2 or 3 at each synaptic step (Rolls, 2000). From this known organization, we predicted that a bigram detector, tuned for instance to bigram "BA," should not be able to respond identically whatever the position of the "B" to the left of an "A" (Dehaene et al., 2005). Rather, letter spacing should matter, and a neuron responsive to "BA" should only be able to cumulate input from hierarchically earlier detectors for letters "B" and "A" when their receptive fields are close enough. Hence, a 4-letter word with widely spaced letters should not be treated as a single visual object but as a series of 4 distinct items whose identification requires serial attention. Accordingly, patients with an impaired control of attentional scanning are unable to read words with spaced letters while they are flawless with normal words (Vinckier et al., 2006). On this account, the reason why spacing letters impairs word reading is not simply because such format is unusual but mainly because letter detectors whose receptive fields are too far apart cannot converge on higher level detectors.

Is it possible to predict the critical letter spacing threshold above which reading should be disrupted? Given the increase of receptive fields in IT cortex by a factor of about 2.5 from one neural level to the next (Rolls, 2000), the LCD model proposes that bigram detectors integrate letter information over a range of 2–3 letter positions (Dehaene et al., 2005). They should, therefore, fail to detect their preferred letter pairs whenever the component letters are separated by a blank space too large to allow two letters to fall within its receptive field. For a hypothetical receptive field of 3 letter positions, a spacing of two letter widths should be sufficient to induce a breakdown of parallel reading, while a spacing of one letter width should not have the same impact.

In two previous studies, we examined the impact on reading performance and brain activation of various modes of word degradation, including letter spacing (Cohen et al., 2008; Vinckier et al., 2006). As expected, a threshold of about 2 blank spaces was indeed necessary for reading performance to deteriorate and for a length effect to emerge, both in normal subjects studied with fMRI and in a patient with parietal damage. However, the main goal of those studies was to demonstrate the intervention of parietal areas whenever stimulus degradation requires the serial deployment of attention to word parts rather than to investigate the role of letter spacing per se. As a consequence, we did not establish beyond doubt whether letter spacing was really the critical feature or whether performance degradation resulted from correlated parameters such as overall stimulus size or eccentricity.

In the present paper, our aim is to disentangle the intrinsic impact of letter spacing on reading from the contribution of potential artifacts and to determine the minimum spacing that is required to deteriorate reading performance and induce a length effect. Our expectations are that letter spacing should interfere with reading independently from correlated parameters and that the spacing threshold should be of about 2 blank spaces.

Experiment 1

The goal of this first experiment was to establish whether introducing blank space between letters deteriorates reading performance and yields a word length effect and to estimate the value of the critical spacing threshold. Spacing letters has the inescapable consequence of increasing the physical size of letter strings. In order to tease apart the role of spacing and of size, we, therefore, used a control condition in which the size of stimuli was increased by an equal amount by using larger fonts, while keeping a normal spacing between letters.

Note that our aim was to study the early, visual, component of word reading. Therefore, rather than asking subjects to read words aloud, we used a lexical decision task. This task allows for a precise measurement of response latencies to printed words, requiring full encoding of the stimuli, while avoiding several sources of variability associated to oral output (Ferrand & New, 2003). Naturally, we expected (and verified in Experiment 2b) that our conclusions do apply to more natural reading conditions.

Methods

Participants

Twelve right-handed native French speakers participated in this experiment (7 men and 5 women; mean age 24 years). All had normal or corrected-to-normal vision and were naive about the aims of the experiment.

Materials

Three sets of 50 four-, six-, and eight-letter high-frequency words were created (frequency of 20–50 per million; New, Pallier, Brysbaert, & Ferrand, 2004). The three sets were matched for word frequency (P = 0.28), letter frequency (P = 0.46), and bigram frequency (P = 0.49; Table 1). Three sets of 50 pseudowords were created, matched one by one with words in terms of consonant–vowel structure, both phonologically and orthographically (e.g., MOUTON and DAIRET). The quality of pseudowords as potential French words was checked by three naive native French speakers. Sixty-five percent of pseudowords had real words among their close orthographic neighbors, defined as substitution neighbors (e.g., OTARUE ≫ OTARIE), deletion neighbors (COURAGNE ≫ COURAGE), or addition neighbors (MIER ≫ MIMER). We also checked that all pseudowords shared their first and last letters with at least one familiar French word of the same length (e.g., CITROL and CHEVAL). Targets were presented in uppercase Arial, white on a black background, within the central 10 degrees of the visual field.

Stimuli were presented in two possible modes (Spacing and Font Size), with five possible values of the scaling factor (Figure 1). At scaling 0, Font Size and Spacing condition were identical, consisting of strings of normally spaced 7-pt letters (letter height and maximum width: 0.27°). In the Spacing mode, increasing the scaling factor was achieved by increasing the number of blank spaces (1 to 4) between letters, while keeping letter size constant. In the Font Size mode, increasing the scaling factor was achieved by increasing the size of letters so as to match the length of letter strings at the same scaling factor in the Spacing mode, while keeping a normal spacing of letters.

Procedure

Each trial started with a 690-ms fixation point, which was replaced by a word or a pseudoword, centered on fixation. Subjects were instructed to maintain their gaze on the fixation point all through experimental blocks (there was a break every 150 trials). They were asked to perform a lexical decision task and to respond by pressing a button with their left hand for pseudowords and with their right hand for real words. The target remained visible until subjects responded.

All words and pseudowords were presented once in the Spacing mode and once in the Font Size mode. In each mode, a given item was associated with a randomly selected scaling factor. Stimuli were presented in a different random order to each subject. An additional set of 20 training trials was run before the experimental list.

Results

Error rates and median correct RTs for real words were computed for each subject and each condition and were entered in ANOVAs with 3 within-subject factors (presentation mode, scaling factor, and length in number of letters) and subjects as random factor (Figure 2 and Table 2). Note that only responses to real words were included in the analysis because response times to pseudowords may mostly reflect the failure to contact the lexicon after a

	Letters’ textual frequency	Bigrams’ textual frequency	Morphologically simple
Four letters (first set)	80,036	9474	50/50
Six letters (first set)	81,395	9994	47/50
Eight letters (first set)	81,849	9915	45/50
Four letters (second set)	81,216	9808	150/150
Six letters (second set)	83,710	10,017	142/150

Table 1. Word properties.

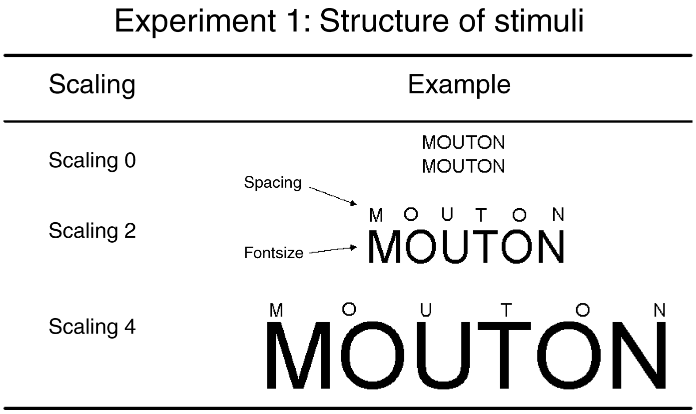

Figure 1. Structure of stimuli for Experiment 1. In order to tease apart the role of letter spacing and of physical stimulus length, the size of letter strings was varied either by separating letters by up to 4 blank spaces or by increasing the size of the font. Stimuli were presented centrally.

given time. Therefore, responses to pseudowords show little influence of low-level visual properties of the stimuli, which are the focus of the current study.

Error rates

All subjects made less than 10% errors. There was no significant effect of length and no interaction involving this factor. There was an interaction of presentation mode and scaling factor ($F(4, 44) = 5.3$; $P = 0.001$). In the Spacing mode, error rate increased with scaling ($F(4, 44) = 7.59$; $P < 0.001$), while it did not differ across scaling values in the Font Size condition ($P > 0.1$).

Response times

Reaction time data showed essentially the same pattern as error rates, plus an impact of word length. There was an interaction of presentation mode, scaling, and length ($F(8, 88) = 7.34$; $P < 0.001$), and the two modes were, therefore, analyzed separately.

In the Font Size mode, there was no effect of length or scaling and no interaction of those factors. In the Spacing condition, RTs increased with larger spacing ($F(4, 44) = 54.2$; $P < 0.001$) and with words of increasing length ($F(2, 22) = 28$; $P < 0.001$). There was an interaction of those 2 factors ($F(8, 88) = 13.2$; $P < 0.001$), as the effect of word length emerged and increased only for scaling factors of 2 or more. Pairwise comparisons showed that latencies increased between consecutive scaling values from 1 to 4 (1 to 2: 50 ms; $F(1, 11) = 7.1$; $P = 0.004$; 2 to 3: 43 ms; $F(1, 11) = 16$; $P < 0.001$; 3 to 4: 80 ms; $F(1, 11) = 49$; $P < 0.001$). The difference between scaling 0 and 1 was small (11 ms) and non-significant ($P > 0.1$).

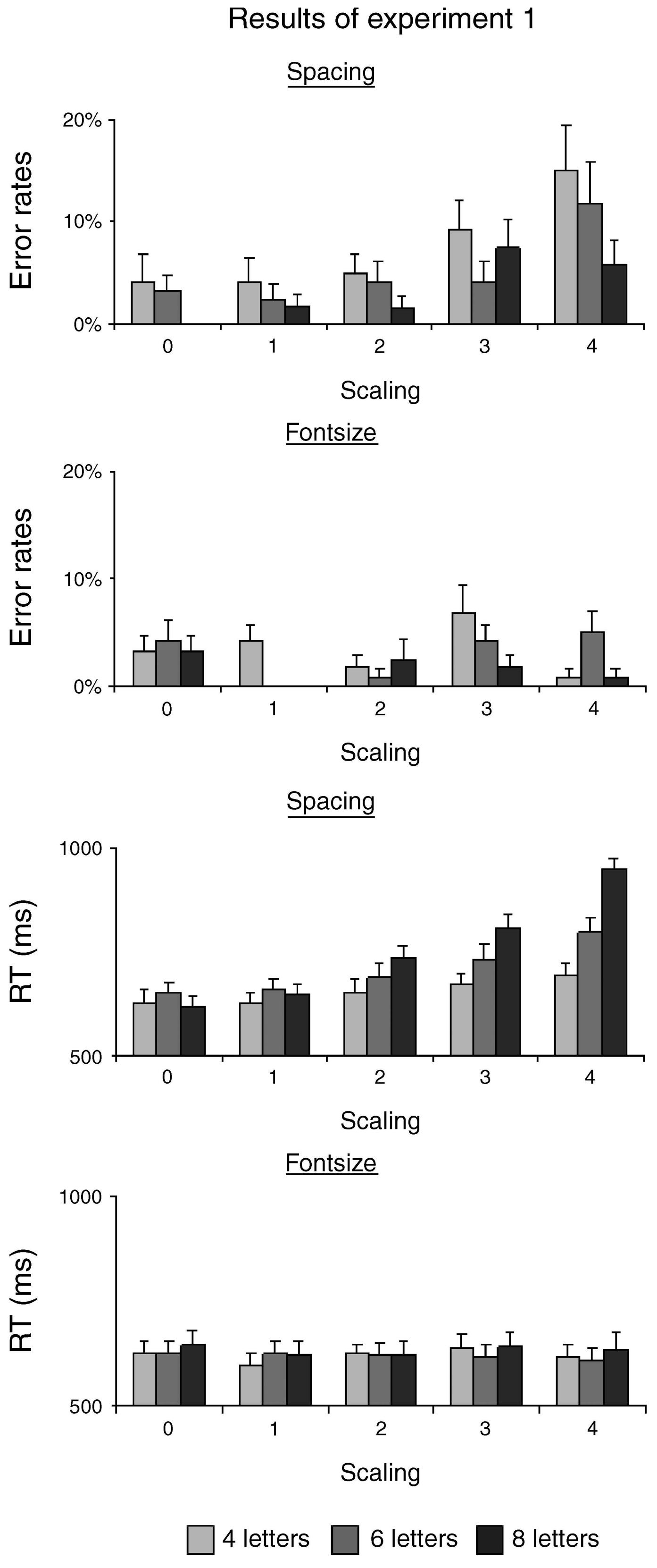

Figure 2. Results of Experiment 1. Error rates and RTs increased whenever letters were separated with at least 2 blank spaces, while they were not affected by increasing font size. Moreover, for spacing values of 2 or more, a word length effect emerged.

Error rate	Spacing condition					
	Scaling value					
No. of letters	0	1	2	3	4	All
4	4.2%	4.2%	5.0%	9.2%	15.0%	7.5%
6	3.3%	2.5%	4.2%	4.2%	11.7%	5.2%
8	0.0%	1.7%	1.6%	7.5%	5.8%	3.3%
All	2.5%	2.8%	3.6%	7.0%	10.8%	5.3%

Error rate	Font size condition					
	Scaling value					
No. of letters	0	1	2	3	4	All
4	3.3%	4.2%	1.7%	6.7%	0.8%	3.3%
6	4.2%	0.0%	0.8%	4.2%	5.0%	2.8%
8	3.3%	0.0%	2.5%	1.7%	0.8%	1.7%
All	3.6%	1.4%	1.7%	4.2%	2.2%	2.6%

Mean RT	Spacing condition					
	Scaling value					
No. of letters	0	1	2	3	4	All
4	625	623	650	668	695	652
6	652	659	687	730	796	705
8	615	643	737	806	952	751
All	630	642	691	735	814	703

Mean RT	Font size condition					
	Scaling value					
No. of letters	0	1	2	3	4	All
4	626	597	624	640	617	621
6	629	624	623	619	613	621
8	647	622	623	641	636	634
All	634	614	623	633	622	625

Table 2. Results of Experiment 1.

Analyses restricted to each value of scaling showed that the length effect was significant for values 2 to 4 (2: $F(2, 22) = 12.8$; $P < 0.001$; 3: $F(2, 22) = 10.2$; $P < 0.001$; 4: $F(2, 22) = 47.3$; $P < 0.001$). The size of the length effect increased with scaling ($R = 0.82$, $P < 0.001$; Figure 2).

Discussion

In summary, error rates and RTs increased whenever letters were separated with at least 2 blank spaces, while they were not affected by increasing font size. Moreover, for spacing values of 2 or more, a word length effect emerged. This length effect then increased with wider spacing. Most importantly, simply manipulating the size of letters while keeping them normally spaced had no impact on performance, demonstrating that the effect of spacing was not an artifact related to word size.

Those results are consistent with the general prediction that spacing letters should impair reading performance and induce an effect of length. Both performance degradation and the length effect appeared for the same value of spacing, supporting the idea of a common underlying mechanism. Furthermore, the value of this spacing threshold, around two, fitted our quantitative expectations based on the physiology of the ventral visual cortex.

However, an alternative account of the increasing reading difficulty associated with spacing should be considered. A consequence of introducing blank space between letters is to move letters farther away from fixation on average. As visual acuity decreases with eccentricity, performance degradation could be due to peripheral viewing and not to spacing per se. This loss of acuity would not affect performance in the Font Size condition, because the shift to the periphery is compensated by the associated increase in the size of letters. It should be noted that, even if performance degradation partly reflects the eccentricity of the outermost letter, such an effect may, in principle, coexist with an actual effect of letter spacing. Indeed, the results observed in the Spacing condition provide some indications that RTs were not fully determined by the location of the outermost letters. Thus, RTs were significantly shorter for 8-letter stimuli at scaling 1 (643 ms) than for (a) 4-letter stimuli at scaling 3 (668 ms), (b) 4-letter stimuli at scaling 4 (695 ms), and (c) 6-letter stimuli at scaling 2 (687 ms; all $P < 0.05$). This was true even though the location of the outermost letter was approximately 7.5 letter for all of those conditions, showing that performance did not depend only on maximum eccentricity.

Those arguments are, however, not sufficient to disentangle the contributions of eccentricity and spacing, which was the aim of Experiment 2a. As a point of method, note that in Experiment 1, stimuli remained visible until subjects responded, and the occurrence of eye movements could not be excluded in spite of task instructions. In the following experiments, stimuli were briefly flashed, so as to prevent eye movements.

Experiment 2a

The goal of Experiment 2a was to determine whether the degradation of reading performance induced by spacing letters resulted solely from the average shift of letters to the periphery of the visual field induced by spacing or from this effect plus a specific effect of spacing. To this end, eccentricity and spacing were manipulated so as to yield contrasting predictions. Schematically, for each value of letter spacing, performance was compared between spaced stimuli and stimuli with contiguous letters but with a larger average letter eccentricity. If performance is worse for spaced stimuli than for the corresponding displaced stimuli, it would imply that spacing has a deleterious effect of its own, above and beyond the effect of eccentricity. Because the impact of eccentricity on reading performance differs across the two visual hemifields, with a more severe degradation with increasing eccentricity in the left than in the right hemifield (Ellis, 2004), the manipulation of eccentricity and spacing was fully crossed with the hemifield in which words were presented.

Methods

Participants

Eighteen subjects participated in this experiment (11 men and 7 women, mean age 23 years), obeying the same criteria as in Experiment 1.

Materials

We used a subset of the words and pseudowords from Experiment 1, including only the 4- and 6-letter stimuli. Targets were presented in uppercase Arial 7-pt font, white on a black background, within the central 10 degrees of the visual field.

Stimuli were presented in two possible modes (Spacing and Displacement), with five possible values of the scaling factor, in either the left or the right visual hemifield (Figure 3). At scaling 0, Spacing and Displacement conditions were identical, consisting of strings of contiguous 7-pt letters (letter height and maximum width: 0.27°) displayed in one hemifield. In the Spacing mode, increasing the scaling factor was achieved by increasing the width of the blank space between letters (0.6, 1.2, 1.8, and 2.4 spaces). In the Displacement mode, increasing the scaling factor was achieved by shifting stimuli away from fixation, while keeping a normal spacing of letters.

For any given value of the scaling factor, the lateral edge of all targets was aligned with the lateral edge of 6-letter words in the spacing mode. Therefore, all targets were justified at a same maximal eccentricity (Figure 3).

Procedure

Each trial started with a 690-ms fixation point. It was replaced by a word or a pseudoword that remained visible for 170 ms, in order to avoid saccades and foveation of stimuli. Subjects were instructed to perform a lexical decision task and to respond by pressing a button with their left hand to pseudowords and with their right hand to real words. The next trial was triggered by the response.

All words and pseudowords were presented once in the Spacing mode and once in the Displacement mode. In each mode, a given word was associated with a randomly selected scaling factor and with a randomly selected hemifield. Stimuli were presented in a different random order to each subject. An additional set of 60 training trials was run before the experimental list.

Results

Error rates and median correct RTs for real words were computed for each subject and each condition and were entered in ANOVAs with 4 within-subject factors (number of letters, scaling factor, presentation mode, hemifield) and subjects as random factor. There was a significant interaction of hemifield, mode, scaling, and length for both error rates ($F(4, 68) = 2.6$; $P = 0.042$) and response times ($F(4, 60) = 4.33$; $P = 0.004$), and the results were analyzed separately for the two hemifields (Figure 4 and Table 3).

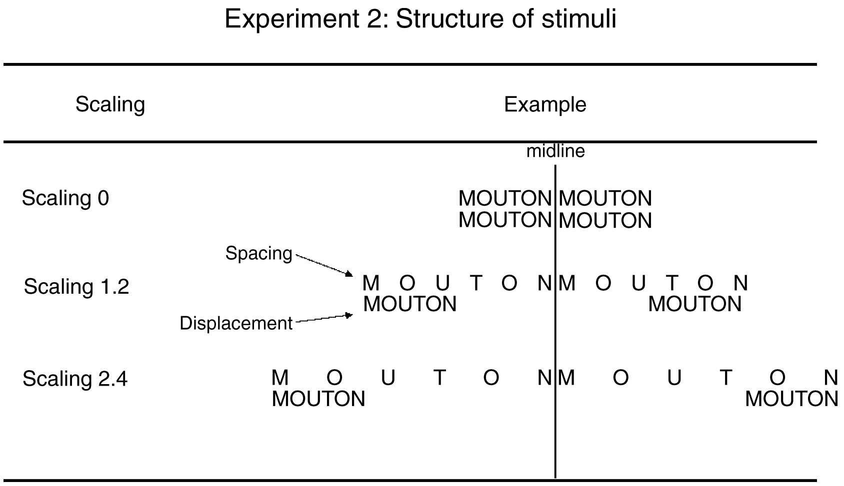

Figure 3. Structure of stimuli for Experiment 2a. In order to tease apart the role of letter spacing and of eccentricity, stimuli were manipulated either by separating letters by up to 2.4 blank spaces or by displacing stimuli toward the periphery of the visual field. Stimuli were flashed in the left or right hemifield.

Error rates

There was an overall right-hemifield advantage as commonly reported in reading experiments ($F(1, 17) = 6.8$; $P = 0.018$).

Left hemifield: There was a significant interaction of scaling and mode ($F(4, 68) = 2.6$; $P = 0.04$). Error rate increased with scaling ($F(4, 68) = 11.8$; $P < 0.001$) in both the Spacing and Displacement modes ($F(4, 68) = 10.8$; $P < 0.001$ and $F(4, 68) = 4.5$; $P = 0.003$, respectively). For scaling value of 0.6, there was a marginally significant effect of mode ($F(1, 17) = 4.0$; $P = 0.061$), with somewhat higher error rates in the Displacement condition. Only for

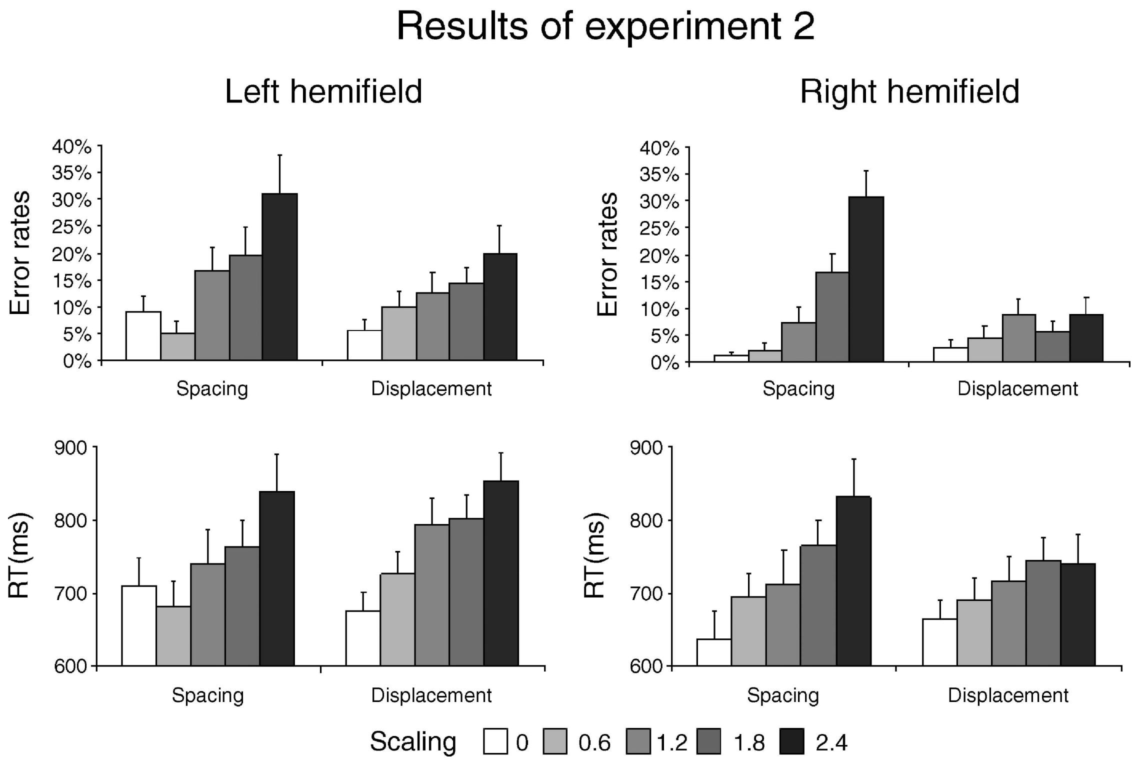

Figure 4. Results of Experiment 2a. Both letter spacing and stimulus displacement had a deleterious impact on reading performance. The effect of spacing was mostly visible above a value of 1.8 spaces. Beyond this threshold, reading was more difficult for spaced than for the corresponding displaced stimuli, although in displaced stimuli, letters were, on average, more eccentric than in spaced stimuli.

	Left hemifield: Spacing condition					
Error rate	Scaling value					
No. of letters	0	0.6	1.2	1.8	2.4	All
4	7.8%	2.2%	18.7%	13.3%	24.4%	13.3%
6	10.0%	7.8%	14.3%	25.6%	37.6%	19.1%
All	8.9%	5.0%	16.5%	19.5%	31.0%	16.2%
Error rate	Left hemifield: Displacement condition					
	Scaling value					
No. of letters	0	0.6	1.2	1.8	2.4	
4	3.3%	7.8%	13.3%	17.4%	15.6%	11.5%
6	7.8%	11.9%	11.9%	10.9%	24.4%	13.4%
All	5.6%	9.9%	12.6%	14.2%	20.0%	12.4%
Error rate	Right hemifield: Spacing condition					
	Scaling value					
No. of letters	0	0.6	1.2	1.8	2.4	All
4	2.2%	4.4%	10.0%	21.1%	38.9%	15.3%
6	0.0%	0.0%	4.4%	12.2%	22.2%	7.8%
All	1.1%	2.2%	7.2%	16.7%	30.6%	11.5%
Error rate	Right hemifield: Displacement condition					
	Scaling value					
No. of letters	0	0.6	1.2	1.8	2.4	
4	5.4%	4.4%	10.0%	6.7%	12.2%	7.7%
6	0.0%	4.4%	7.8%	4.4%	5.6%	4.4%
All	2.7%	4.4%	8.9%	5.6%	8.9%	6.1%
Mean RT	Left hemifield: Spacing condition					
	Scaling value					
No. of letters	0	0.6	1.2	1.8	2.4	All
4	697	671	705	735	749	712
6	724	694	773	791	926	782
All	710	683	739	763	838	747
Mean RT	Left hemifield: Displacement condition					
	Scaling value					
No. of letters	0	0.6	1.2	1.8	2.4	
4	662	708	781	777	882	762
6	690	744	808	829	823	779
All	676	726	794	803	853	770
Mean RT	Right hemifield: Spacing condition					
	Scaling value					
No. of letters	0	0.6	1.2	1.8	2.4	All
4	631	721	725	786	862	745
6	644	668	697	744	801	711
All	638	695	711	765	831	728
Mean RT	Right hemifield: Displacement condition					
	Scaling value					
No. of letters	0	0.6	1.2	1.8	2.4	All
4	679	708	728	773	739	725
6	651	674	704	717	743	697
All	665	691	716	745	741	711

Table 3. Results of Experiment 2a.

scaling value of 2.4, error rates were higher in the Spacing mode than in the Displacement mode ($F(1, 17) = 5.0$; $P = 0.04$).

Right hemifield: There was a significant interaction of scaling and mode ($F(4, 68) = 23.7$; $P < 0.001$). Error rate increased with scaling in both the Spacing and Displacement modes ($F(4, 68) = 35.7$; $P < 0.001$ and $F(4, 68) = 3.7$; $P = 0.009$, respectively). In Spacing mode, there was a steep jump between scaling 1.2 and 1.8. In contrast, there was no significant effect of scaling in the Displacement mode. Only for scaling values of 1.8 and 2.4, error rates were higher in the Spacing mode than in the Displacement mode ($F(1, 17) = 23.9$; $P < 0.01$ and $F(1, 17) = 44.2$; $P < 0.001$, for scaling 1.8 and 2.4, respectively).

Response times

Two subjects were removed from this analysis because they produced no correct response to real words in the most difficult condition (Spacing mode, left hemifield, scaling 2.4). The pattern was similar to that observed with error rates, showing different profiles in the two hemifields (Figure 4). There was again an overall right-hemifield advantage ($F(1, 15) = 18.7$; $P < 0.001$).

Left hemifield: There was an interaction of mode and scaling ($F(4, 60) = 3.4$; $P = 0.015$). RTs increased with scaling in both the Spacing and Displacement modes ($F(4, 60) = 14.8$; $P < 0.001$ and $F(4, 60) = 21.9$; $P < 0.001$, respectively). However, for the lower two scaling values (0.6 and 1.2), RTs were slower in the Displacement mode ($F(1, 15) = 6.4$; $P = 0.02$ and $F(1, 15) = 5.4$; $P = 0.04$, respectively). For the highest scaling values, there was no effect of mode.

Right hemifield: There was again an interaction of scaling and mode ($F(4, 60) = 5.0$; $P = 0.001$). RTs increased with scaling in both the Spacing and Displacement modes ($F(4, 60) = 19.5$; $P < 0.001$ and $F(4, 60) = 4.8$; $P = 0.002$, respectively). In the Spacing conditions, the increase was steep, with RTs rising to 830 ms. In the Displacement condition, the increase stopped at about 740 ms, with slower RTs in the Spacing condition than in the Displacement condition only for scaling 2.4 ($F(1, 15) = 14.1$; $P = 0.002$).

Effect of word length

The impact of word length differed across hemifields (interaction length × hemifield for errors: $F(1, 17) = 19.2$; $P < 0.001$; and for RTs: $F(1, 15) = 27.4$; $P < 0.001$). Within each hemifield, word length did not interact with scaling and mode. In the left hemifield, a word length effect was observed in the usual direction, i.e., easier reading of shorter words (errors: $F(1, 17) = 5.71$; $P = 0.03$; RTs: $F(1, 15) = 22.0$; $P < 0.001$). However, in the right hemifield, the length effect was reversed, i.e., easier reading of longer words (errors: $F(1, 17) = 19.7$; $P < 0.001$; RTs: $F(1, 15) = 10.7$; $P = 0.005$).

Discussion

The main goal of Experiment 2a was to determine whether the deleterious effect of letter spacing on reading performance was an artifactual consequence of letter eccentricity, which in Experiment 1 was positively correlated with spacing. Here, by using a condition with high eccentricity but normal letter spacing, we had the opportunity to disentangle the contribution of those two parameters. In brief, we observed that both modes of stimulus degradation had a deleterious and independent impact on reading performance. First, unsurprisingly, reading performance decreased when stimuli were shifted away from fixation. This effect was more important in the left hemifield than in right hemifield, in agreement with previous evidence that the optimal reading area extends farther in the right hemifield than in the left hemifield (Nazir, Jacobs, & O'Regan, 1998; Rayner & Bertera, 1979). Second, we observed an effect of spacing, mostly visible above a threshold value of about 1.8 spaces. This effect of spacing was comparable in the two hemifields. Note also that there was excellent quantitative agreement between Experiments 1 and 2a as to the critical spacing threshold of about 2 blank spaces or a bit less between consecutive letters.

In this experiment, the crucial issue was the relative impact of the two modes of degradation. We found that beyond the spacing threshold already identified in Experiment 1, reading was more difficult for spaced than for the corresponding displaced stimuli, although in displaced stimuli, letters were, on average, more eccentric than in spaced stimuli and, being contiguous, affected by more severe crowding effects. This result clearly demonstrates that the impact of spacing on reading performance cannot be reduced to an artifact of eccentricity.

This conclusion is clear-cut for right-hemifield stimuli, with converging analyses of error rates and RTs. For left-hemifield stimuli, displacement to the periphery had a more severely disruptive effect, possibly due to the greater eccentricity of the initial letters, which are the most informative letters for word recognition (O'Regan, Levy-Schoen, Pynte, & Brugaillere, 1984). The contrast between the Displacement and Spacing conditions was, therefore, less marked and significant only with error rates. However, even in the left hemifield, the mere fact that Spacing was more rather than less disruptive than Displacement supports our main conclusion.

In the current experiment, we observed a clear-cut inverse word length effect in the right hemifield. This unusual response pattern was probably linked to the unusual word display, with short and long words aligned by their peripheral edge rather by their initial letter as in most studies. At first sight, this inverse effect could be due to the fact that in the right hemifield the initial letters were much more eccentric and, therefore, more difficult to identify, in short than in long words, particularly in the spacing mode. However, this simple hypothesis may not be sufficient. It would predict that in the right visual field, for a given number of letters and a given scaling (i.e., for a given maximum eccentricity), words with spaced letters (i.e., in Spacing mode) should be easier to read than words with contiguous letters (i.e., in Displacement mode), just because the latter start farther away from fixation. Naturally, this prediction should hold for spacing values below the putative threshold of 2 spaces, i.e., when spacing does not have a deleterious effect by itself. Actually, as shown in Table 3, those two types of trials did not differ: for 4- and 6-letter words, at scaling 0.6 and 1.2, latencies did not differ between the Spacing and Displacement modes. This seems to disconfirm the idea that the inverse length effect would reflect the eccentricity of the first letter. However, the spaced and displaced words that we just compared differed in several respects in addition to the eccentricity of their first letter: they were not comparable in spacing, crowding, mean eccentricity, or physical width. Actually, the present experiment by itself was inappropriate to study and understand the influence of word length, which was partially confounded with other visual factors, which may interact in a way difficult to predict quantitatively. Interestingly, the SERIOL model of word reading incorporates a lexical access component that would generate an inverse length effect (Whitney & Lavidor, 2004). Detectors for longer words would "settle" faster in the lexicon, due to their smaller number of competitors. This inverse length effect would be cancelled out by a serial letter encoding component that would take longer time for longer words. Schematically, the combination of those two influences would not operate identically in the two hemifields, explaining the usual pattern of asymmetry, i.e., a length effect restricted to the left hemifield. However, although it predicts an asymmetry in length effect, it is not clear which of this model's parameters should be modified to yield the present pattern of results. In Experiment 3 below, we will provide additional evidence that eccentricity of the initial letter is an important determinant of length effects.

Experiment 1 demonstrated that spacing letters by at least two blank spaces slows down reading, with the simultaneous emergence of a word length effect. We took this result as support to our hypothesis that whenever the interval between letters is larger than about 2 spaces, parallel reading collapses and readers resort to serial reading. Experiment 2a showed that the impact of spacing on reading performance was not an artifact of letter eccentricity, a parameter that, in centrally presented words, is correlated with spacing.

There is, however, an alternative interpretation to the difference between spaced and displaced words as observed in Experiments 1 and 2a. This alternative rests on two plausible assumptions. The first assumption is that words are read serially whenever letters are remote from fixation, even for normally spaced letters (at least when eccentricity is not compensated for by an increase in font size as in Experiment 1). The second assumption is that serial letter scanning takes more time and is more error prone for physically longer words, i.e., when attention movements must cover a larger expense of space. Then, at high scaling values in Experiment 2a, both spaced and displaced words would be read serially, but spaced words would be more difficult due to their larger physical size (and not to spacing per se). Note that according to this hypothesis, the effect of spacing would be an artifact of physical size distinct from the one considered in Experiment 1. Experiment 3 was aimed at assessing this alternative interpretation of Experiment 2a. However, before presenting and assessing this alternative account, we wanted to check whether, beyond lexical decision, the critical results of Experiment 2a generalized to more natural reading conditions.

Experiment 2b

The goal of Experiment 2b was to replicate the results of Experiment 2a using a more natural word naming task. As reading aloud should involve the same input processing as lexical decision, we expected to observe the same effects of spacing as before.

Methods

Participants

The subjects were the same as in Experiment 2a. They participated in Experiment 2b just after completion of Experiment 2a.

Materials

A set of 5-letter high-frequency words was created (frequency of 20–50 per million; New et al., 2004).

Procedure

Targets were presented as in Experiment 2a. For any given value of the scaling factor, the lateral edge of all targets was aligned with the lateral edge of 6-letter words in the spacing mode. Subjects were instructed to name stimuli aloud. The next trial was triggered by the response.

All words were presented once in the Spacing mode and once in the Displacement mode. In each mode, a given word was associated with a randomly selected scaling factor and with a randomly selected hemifield. Stimuli were presented in a different random order to each subject.

Results

Error rates were computed for each subject and each condition and were entered in ANOVAs with 3 within-subject factors (scaling factor, presentation mode, hemifield) and subjects as random factor. Note that even in the most difficult condition (Spacing mode at scaling 2.4 in the left hemifield), subjects correctly identified more than 70% of the words (Table 4). There was a significant interaction of mode and scaling ($F(4, 68) = 4.0$; $P = 0.005$) and a main effect of hemifield ($F(1, 17) = 6.24$; $P = 0.02$), with a right-hemifield advantage.

Error rates increased with scaling in both the Spacing and Displacement modes ($F(4, 68) = 12.74$; $P < 0.001$ and $F(4, 68) = 2.9$; $P = 0.03$, respectively). In the Spacing conditions, the increase was steep, with error rates rising to 28%. In the Displacement condition, the increase stopped at about 17%, with higher error rates in the Spacing condition than in the Displacement condition only for scaling 2.4 ($F(1, 17) = 14.3$; $P = 0.002$).

Discussion

The main goal of Experiment 2b was to replicate the results of Experiment 2a with a more natural reading task. As expected, we observed the same effect of spacing than in Experiment 2a. Beyond the spacing threshold identified in Experiments 1 and 2a, reading was more difficult for spaced than for the corresponding displaced stimuli, although in displaced stimuli, letters were, on average, more eccentric than in spaced stimuli. We now turn to the assessment of an alternative account of our results, as presented in the conclusion of Experiment 2a.

Experiment 3

The goal of Experiment 3 was to determine whether, in laterally presented words, reading performance was dependent on the physical length of words rather than on the spacing of letters. To this end, spacing and physical size were manipulated so as to yield contrasting predictions. Schematically, we compared 4-, 6-, and 8-letter words with an identical physical size, which was achieved by a wider spacing of letters in 4- than in 6- than in 8-letter words. If for peripheral words performance depends only on physical size, performance should be comparable irrespective of the number of letters. Conversely, if spacing per se has a critical impact on reading, then words with the smaller spacing, i.e., 8-letter words, should be paradoxically easier to read despite their larger number of letters.

Methods

Participants

Twelve subjects participated in this experiment (5 men and 7 women, mean age 22 years), obeying the same criteria as in previous experiments.

	Left hemifield: Spacing condition					
Error rate	Scaling value					
No. of letters	0	0.6	1.2	1.8	2.4	All
5	6.1%	6.9%	7.70%	26.9%	28.0%	17.0%
Error rate	Left hemifield: Displacement condition					
	Scaling value					
No. of letters	0	0.6	1.2	1.8	2.4	All
5	5.6%	5.6%	11.1%	16.7%	16.7%	11.1%
Error rate	Right hemifield: Spacing condition					
	Scaling value					
No. of letters	0	0.6	1.2	1.8	2.4	All
5	5.6%	0.0%	6.1%	13.6%	25.0%	10.1%
Error rate	Right hemifield: Displacement condition					
	Scaling value					
No. of letters	0	0.6	1.2	1.8	2.4	All
5	4.4%	1.4%	3.6%	7.8%	5.6%	4.6%

Table 4. Results of Experiment 2b.

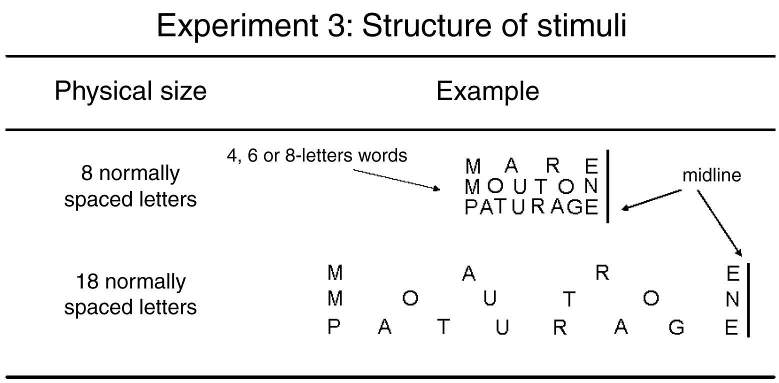

Figure 5. Structure of stimuli for Experiment 3. In order to tease apart, in laterally presented words, the role of letter spacing and of physical length, we compared 4-, 6-, and 8-letter words with an identical physical size, which was achieved by a wider spacing in words with fewer letters. This was done while independently varying the physical length of stimuli. Stimuli were flashed in the left or right hemifield.

Materials

We used a subset of the stimuli of Experiment 1, consisting in 40 items from each of the 6 lists of 50 four-, six-, or eight-letter words and pseudowords. Targets were presented in uppercase Arial 7-pt font, white on a black background, within the central 10 degrees of the visual field. Four values of physical size were used, corresponding to the dimension of normally printed 8-, 12-, 15-, and 18-letter words. The spacing between letters was adjusted in order for 4-, 6-, and 8-letter stimuli to fit exactly in each physical size. Therefore, for each physical size, 4-, 6-, and 8-letter stimuli occupied exactly the same display area (Figure 5). This procedure resulted in spacing values ranging from 0 (for 8-letter words with a physical size of 8) to 4.7 spaces (for 4-letter words with a physical size of 18; Table 5). Stimuli were presented in the left or right hemifield, like in the Spacing mode from Experiment 2a, i.e., adjacent to the fixation point.

Procedure

The task and trial structure were the same as in Experiment 2a. All words and pseudowords were presented once in the left hemifield and once in the right hemifield. In each hemifield, a given word was associated to a randomly selected physical size. Stimuli were presented in a different random order to each subject. An additional set of 48 training trials was run before the experimental list.

Results

Error rates and median correct RTs for real words were computed for each subject and each condition and were entered in ANOVAs with 3 within-subject factors (number of letters, physical size, hemifield) and subjects as random factor (Figure 6 and Table 6).

Error rates

There was a right-hemifield advantage ($F(1, 11) = 14.0$; $P = 0.003$) and an interaction of number of letters and physical size ($F(6, 66) = 2.9$; $P = 0.02$). For small physical size (8 letters), there was no effect of number of letters ($P > 0.1$), whereas there was an effect of number of letters for all larger physical sizes ($F(2, 22) = 4.2$, $P = 0.03$; $F(2, 22) = 3.7$ $P = 0.04$; and $F(2, 22) = 6.2$; $P = 0.07$ for physical sizes 12, 15, and 18, respectively).

Responses times

One subject was removed from this analysis because he produced no correct response in two of the most difficult left-hemifield conditions (8 letters, size 15; and 6 letters, size 18). There were no significant interactions. There was the usual right-hemifield advantage ($F(1, 10) = 23.63$; $P < 0.001$). Latencies increased with larger physical size ($F(3, 30) = 23.1$; $P < 0.001$). Finally; latencies were slower for words with fewer letters ($F(2, 20) = 9.67$; $P = 0.0012$; similar to the pattern of errors, there was a tendency for this length effect to be larger in the right hemifield than in the left hemifield; interaction $F(2, 20) = 2.7$; $P = 0.09$).

Length effect

In Experiment 2a, we observed a reversal of the word length effect across hemifields. In the right hemifield,

No. of spaces	Physical size (in number of normally spaced letters)			
No. of letters	8	12	15	18
4	1.3	2.7	3.7	4.7
6	0.4	1.2	1.8	2.4
8	0.0	0.6	1.0	1.4

Table 5. Structure of stimuli of Experiment 3.

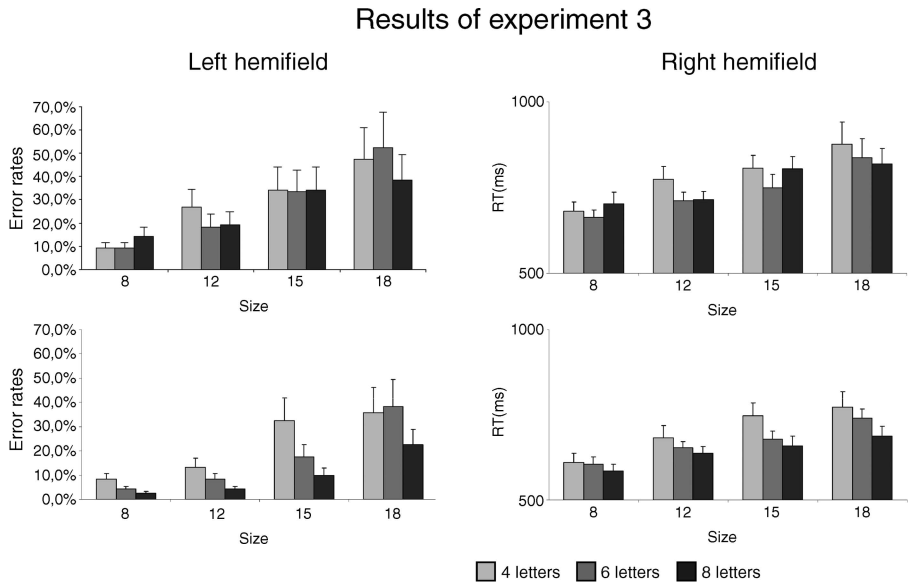

Figure 6. Results of Experiment 3. Overall reading performance deteriorated when words were physically longer. The critical result is that, for any given physical size, performance was worst for words comprising fewer letters, i.e., for larger values of spacing, confirming that the impact of spacing cannot be reduced to the effect of physical length.

contrary to the usual pattern, shorter words were more difficult to read than longer words. We hypothesized that this was due to the fact that in the right hemifield, the initial letter was more peripheral for short than for long words. If this account is correct, words displayed in the RVF with their initial letter at a constant eccentricity should show the usual length effect. The present experiment gave us an opportunity to test precisely this situation. We compared responses to 4-letter words fitted in the physical size of 12 vs. 6-letter words fitted in the physical size of 18. Those two types of stimuli had approximately the same space between letters (2.7 and 2.4 spaces, respectively). In the right visual field, six-letter words yielded higher error rates ($F(1, 11) = 16.8$; $P = 0.0018$) and marginally longer latencies ($F(1, 10) = 4.23$; $P = 0.067$) than four-letter words. Conversely, when restricting the analysis to right-hemifield stimuli with a spacing smaller than 2 spaces, there was no significant effect of word length ($P > 0.1$). In summary, whenever the eccentricity of the first letter is kept constant, there is no inverse length effect in the right hemifield.

Discussion

Overall, reading performance deteriorated when words were physically larger. Naturally, for words of a given number of letters, physical size is proportional to letter spacing, and the role of the two parameters cannot be dissociated. However, the critical result is that, for any given physical size, performance deteriorated for words comprising fewer letters (an inverse length effect), i.e., for larger values of spacing. It is now possible to answer the question that motivated Experiment 3. Even for laterally presented words, performance depends critically on spacing, an effect that cannot be reduced to the effect of physical size.

Note that there was a tendency for the inverse length effect to be larger in the right hemifield than in the left hemifield. A natural account of this asymmetry is that the inverse length effect was partially cancelled by the usual length effect prevailing in the LVF. It is now safe to conclude from the above experiments that reading performance deteriorates whenever a critical of space is introduced between letters and that this effect cannot be reduced to artifacts of physical size or eccentricity. However, an important point still needs clarification. According to the hypotheses presented in the Introduction section, this threshold should scale with letter size. Thus, it should be about 2 spaces, irrespective of font size, rather than defined by some fixed angular value. However, the first three experiments used the same font size (letter height and maximum width: 0.27°) for all stimuli with spaced letters. The aim of the next experiment is to assess the value of the spacing threshold with other font sizes.

Experiment 4

In order to study the interaction of spacing with absolute font size, we designed a fully crossed experiment combining five degrees of spacing with three font sizes.

	Left hemifield				
Error rate	Physical size (in number of normally spaced letters)				
No. of letters	8	12	15	18	All
4	9.2%	26.7%	34.2%	47.5%	29.4%
6	9.2%	18.3%	33.3%	52.5%	28.3%
8	14.2%	19.2%	34.2%	38.3%	26.5%
All	10.9%	21.4%	33.9%	46.1%	28.1%
Error rate	Right hemifield				
	Physical size (in number of normally spaced letters)				
No. of letters	8	12	15	18	All
4	8.3%	13.3%	32.5%	35.8%	22.5%
6	4.2%	8.3%	17.5%	38.3%	17.1%
8	2.5%	4.2%	10.0%	22.5%	9.8%
All	5.0%	8.6%	20.0%	32.2%	16.5%
Mean RT	Left hemifield				
	Physical size (in number of normally spaced letters)				
No. of letters	8	12	15	18	All
4	681	775	806	876	785
6	663	711	749	837	740
8	703	715	805	820	761
All	682	734	787	844	762
Mean RT	Right hemifield				
	Physical size (in number of normally spaced letters)				
No. of letters	8	12	15	18	All
4	611	681	748	772	703
6	604	653	678	741	669
8	585	636	658	688	642
All	600	657	695	734	671

Table 6. Results of Experiment 3.

Methods

Participants

Thirty-three subjects participated in this experiment (16 men and 17 women, mean age 22 years), obeying the same criteria as in Experiment 1.

Materials

Two sets of 150 four- and six-letter high-frequency words were constructed (frequency of 20–50 per million; New et al., 2004). The two sets were matched for word frequency ($P = 0.33$), letter frequency ($P = 0.62$), and bigram frequency ($P = 0.14$; Table 7). Two sets of 150 pseudowords were created, matched one by one with words in terms of phonological and graphemic CVC structure. The quality of pseudowords as possible French words was checked by three naive native French speakers. Seventy-seven percent of pseudowords had real words among their close orthographic neighbors, defined as substitution neighbors (e.g., MOUSON ≫ MOUTON), deletion neighbors (FINIER ≫ FINIR), or addition neighbors (MIER ≫ MIMER). We also checked that all pseudowords shared their first and last letters with at least one familiar French word of the same length (e.g., CITROL and CHEVAL). Targets were presented in uppercase Arial, white on a black background, and were always within the central 10 degrees of the visual field.

We used a fully crossed design with two values of word length (4 and 6 letters), five values of spacing (0, 0.75, 1.5, 2.25, and 3 spaces), and three font sizes (letter height and maximum width: 0.27°, 0.41°, and 0.54°; Figure 7).

Procedure

The task and procedure were the same as in Experiment 1 except that the targets remained visible for 170 ms, in order to avoid eye movements. All words and pseudowords were presented once. Each stimulus was associated to randomly selected spacing value and font size. Stimuli were presented in a different random order to each subject. An additional set of 20 training trials was run before the experimental list.

Results

Error rate and median correct RT were computed for each subject and each condition and were entered in ANOVAs with 3 within-subject factors (font size, spacing, number of letters) and subjects as random factor (Figure 8 and Table 7).

Error rates

All subjects but one (who was removed from the analysis) made less than 15% errors. Error rates increased with wider spacing ($F(4, 124) = 10.7$; $P < 0.001$), decreased with larger number of letters ($F(1, 31) = 53.9$; $P < 0.001$), and marginally decreased with font size ($F(2, 62) = 2.8$; $P = 0.07$). No interaction was significant, notably no interactions involving font size (Figure 7).

Response times

Just like in the Spacing condition of Experiment 1, RTs increased with larger spacing ($F(4, 124) = 39.09$; $P < 0.001$) and with words of increasing length ($F(1, 31) = 9.36$; $P = 0.005$). The effect of spacing was non-linear: There was no difference between spacing values of 0, 0.75, and 1.5 ($F(1, 31) = 0.004$; $P = 0.95$), while RTs increased for values of 1.5, 2.25, and 3 ($F(1, 31) = 13.37$; $P < 0.001$). There was an interaction of length and spacing ($F(4, 124) = 7.12$; $P < 0.001$), as the effect of word length

Error rate	Font size 0.27					
	Scaling value					
No. of letters	0	0.75	1.5	2.25	3	All
4	8.4%	10.0%	10.0%	12.8%	10.9%	10.4%
6	5.3%	4.4%	5.3%	6.6%	11.9%	6.7%
All	6.9%	7.2%	7.7%	9.7%	11.4%	8.6%
Error rate	Font size 0.41					
	Scaling value					
No. of letters	0	0.75	1.5	2.25	3	All
4	5.6%	10.0%	9.1%	10.6%	11.9%	9.4%
6	1.9%	6.2%	4.4%	6.6%	9.1%	5.6%
All	3.8%	8.1%	6.8%	8.6%	10.5%	7.5%
Error rate	Font size 0.54					
	Scaling value					
No. of letters	0	0.75	1.5	2.25	3	All
4	7.8%	7.5%	8.7%	10.6%	10.3%	9.0%
6	1.6%	3.4%	3.4%	7.5%	8.1%	4.8%
All	4.7%	5.5%	6.1%	9.1%	9.2%	6.9%
Mean RT	Font size 0.27					
	Scaling value					
No. of letters	0	0.75	1.5	2.25	3	All
4	579	570	591	602	618	592
6	558	580	598	644	673	611
All	569	575	594	623	645	601
Mean RT	Font size 0.41					
	Scaling value					
No. of letters	0	0.75	1.5	2.25	3	All
4	562	563	588	577	615	581
6	562	559	598	599	653	594
All	562	561	593	588	634	588
Mean RT	Font size 0.54					
	Scaling value					
No. of letters	0	0.75	1.5	2.25	3	All
4	562	562	582	598	617	584
6	551	568	585	620	665	598
All	556	565	584	609	641	591

Table 7. Results of Experiment 4.

emerged and increased only for spacing values of 2.25 and 3 ($P = 0.0028$ and $P = 0.0015$, respectively). Latencies were slightly longer for the smaller font (mean 601 ms) than for the two larger fonts (mean 588 and 591 ms; $F(2, 62) = 4.09$; $P = 0.02$). Crucially, there was no interaction involving font size: As visible in Figure 7, the spacing threshold was always about 2 spaces, irrespective of the absolute size of the font. For each of the 3 font sizes considered separately, the length effect was absent for all spacing values <2 and significant or marginal for spacing values >2.

Discussion

In summary, Experiment 4 replicated the fundamental effect of letter spacing on reading performance, i.e., emergence of a length effect and performance deterioration for spacing of at least 2 spaces. The novel finding, however, is that this pattern prevailed irrespective of absolute font size: the threshold was constant when expressed in terms of number of spaces, while it varied by a factor of 2 in angular size.

Additional analyses with mixed-effects models

We complemented the classical ANOVAs reported above with analyses using linear mixed models, with items and subjects as crossed random factors (Baayen, Davidson, & Bates, 2008). The aim of these additional analyses was to study the respective contributions of correlated factors such as eccentricity, physical width, and spacing. More specifically, they may allow us to discriminate between alternative accounts of the emergence of a length effect: We claimed that there is an interaction of spacing with the number of letters, as a length effect emerges for spacing of about 2 spaces and above. However, alternatively, this interaction could be described as an interaction of spacing with physical width or of spacing with maximum eccentricity. As we did not fully decorrelate number of letters, spacing, and eccentricity/physical width within a single experiment, these alternative

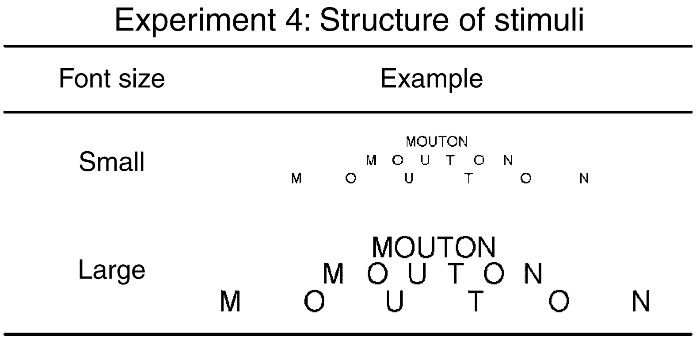

Figure 7. Structure of stimuli for Experiment 4. In order to study the interaction of spacing with absolute font size, we used five degrees of letter spacing with three font sizes in a fully crossed design. Stimuli were presented centrally.

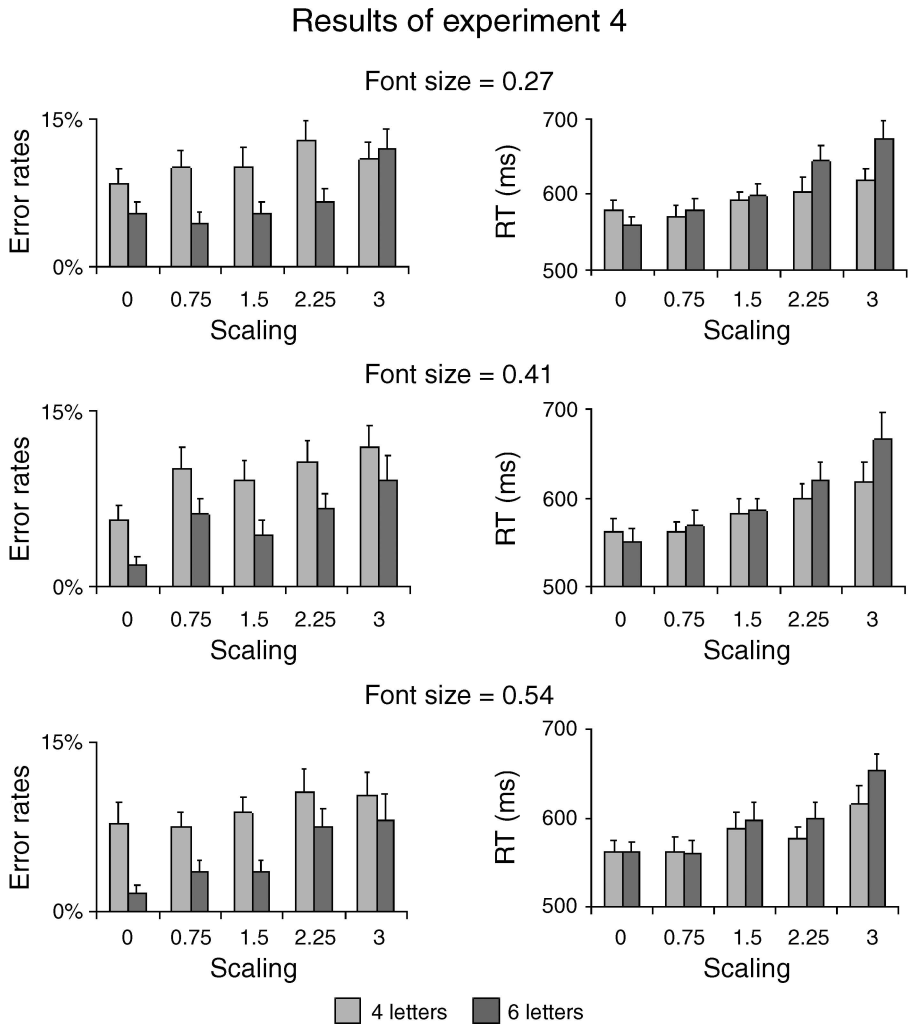

Figure 8. Results of Experiment 4. The emergence of a length effect and performance deterioration for spacing of at least 2 spaces were replicated. This pattern prevailed irrespective of absolute font size: the threshold was constant when expressed in terms of number of spaces, while it varied by a factor of 2 in angular size.

interpretations are difficult to disentangle. However, this issue can be addressed by pooling data across experiments. This is what we did, performing separate analyses for words presented centrally (Experiments 1 and 4) and for words presented laterally, in the RVF or the LVF (Experiments 2 and 3), using mixed-effect regression models, with random intercepts for subjects and items. These analyses were applied to the reaction time data restricted to trials with real words and where the participant's response was correct. RTs were log-transformed prior to analysis to reduce the skewness of the distribution.

Experiments 1 and 4

Data from Experiments 1 and 4 (i.e., centrally presented words) were pooled. Word length and maximal eccentricity, both expressed in number of letters, were included as regressors. We also used a dummy factor "thresholded spacing," which was equal to zero for spacing values <2 and equal to spacing otherwise. We then compared all the possible models with and without interactions. We report here the model with the lowest Akaike Information Criterion (AIC), i.e., with the best trade-off of accuracy and complexity of the model. This model included main effects of word length, maximal eccentricity, and thresholded spacing, plus the interaction of thresholded spacing and length. We applied the Markov Chain Monte Carlo (MCMC) sampling method (with a sample size of 10,000) to obtain P-values for the coefficients (Baayen et al., 2008). Note that the physical width of the strings was not included in the model because it was equal to the maximal eccentricity divided by 2.

In this analysis, RTs increased with thresholded spacing (pMCMC < 0.001) and with maximal eccentricity (pMCMC < 0.001; Table 8). There was no significant

Experiments 1 and 4	Estimate ($*10^{-3}$)	CI ($*10^{-3}$)	*t*-value
No. of letters	−1.3	[−4.2; 1.7]	−0.8
Maximal eccentricity	3.9	[2.5; 5.2]	5.5
Thresholded spacing	8.5	[5.4; 11.5]	5.4
Thresholded spacing: length	4.9	[3.7; 6.1]	8.1
Experiments 2 and 3: Left hemifield			
	Estimate ($*10^{-3}$)	CI ($*10^{-3}$)	*t*-value
Length	5.6	[1.6; 9.5]	2.8
Maximal eccentricity	5.3	[4.2; 6.5]	9.4
Thresholded spacing	8.1	[3.8; 12.4]	3.7
Experiments 2 and 3: Right hemifield			
	Estimate ($*10^{-3}$)	CI ($*10^{-3}$)	*t*-value
Length	−6.7	[−10.5; −2.8]	−3.4
Maximal eccentricity	6.8	[5.8; 7.8]	13
Thresholded spacing	5.2	[1.2; 9.2]	2.6

Table 8. Results of mixed model.

main effect of length (pMCMC > 0.1). However, as in previous analyses, length positively interacted with thresholded spacing (pMCMC < 0.001). Note that we also examined a variant of this model, which included the interaction of thresholded spacing and maximal eccentricity, as this model had a minimally higher AIC. The results of this alternative model were essentially identical, and the additional interaction was not significant (pMCMC > 0.1).[1]

Experiments 2 and 3

Data from Experiments 2 and 3 (i.e., laterally presented words) were pooled. Word length, minimal and maximal eccentricity, and thresholded spacing were included as regressors. We performed this analysis separately for the left and right hemifields. We compared different models (including the full model with all interactions and models with any combination between (i) interaction of thresholded spacing and length, (ii) interaction of maximal eccentricity and length, and (iii) the triple interaction) and kept the model with the lowest AIC in both hemifields. This model included main effects of word length, maximal eccentricity, and thresholded spacing. It did not include the main effect of minimal eccentricity nor any interaction. We applied the Markov Chain Monte Carlo (MCMC) sampling method (with a sample size of 10,000) to obtain *P*-values for the coefficients (Baayen et al., 2008).

In this analysis, RTs increased with maximal eccentricity (pMCMC < 0.001 in the right and in the left hemifield, respectively) and with thresholded spacing (pMCMC $= 0.01$ and pMCMC < 0.001 in the right and left hemifields, respectively; Table 8). In the left hemifield, RTs increased with length (pMCMC $= 0.005$), whereas RTs decreased with length in the right hemifield (pMCMC < 0.001).

Discussion of additional analyses

The results of the mixed-model analyses were consistent with the previous classical analyses. When words were presented centrally (Experiments 1 and 4), a main effect of spacing and an interaction between spacing and number of letters were observed. With lateral presentation (Experiments 2 and 3), the mixed-effects model confirm the main effect of spacing, independently from maximal eccentricity. In summary, we showed that the effects of interest (letter spacing and its interaction with length) are still significant when also modeling the contribution of eccentricity. One should note that although our results broadly fit the LCD model, one prediction was not fulfilled. We have found an interaction of number of letters and spacing for the central presentation, but this interaction was not significant with lateralized presentation, particularly in the right visual field. This lack of interaction is difficult to interpret. Naturally, the model may be inaccurate, and the reading process may not change qualitatively with spacing. For instance, the effect of spacing could be due to increased attentional demand (to group the stimulus as an object) and would not qualitatively change the nature of orthographic analysis. However, such an additional constant attentional cost would not explain the interaction observed with central words, in our data and in other studies (Cohen et al., 2008;

Vinckier et al., 2006). Rather, there may be methodological reasons for not observing this subtle effect, such as the alignment of words by their peripheral edge, with minimal and mean eccentricity larger for short words. Furthermore, only one above-threshold value of spacing was used in lateralized presentation, and this lack of interaction could, thus, be due to a lack of power. This point should be the object of further investigation.

General discussion

Summary of the results

In the present series of experiments, we measured the impact of letter spacing on reading performance and progressively purified this effect from a number of possible confounding variables. In Experiment 1, we established the core phenomenon, namely, that performance deteriorates non-linearly whenever letters are separated by at least 2 blank spaces, with the concomitant emergence of a word length effect. We showed that this effect cannot be reduced to an effect of physical word size, a variable correlated with spacing. Indeed, increasing word size by increasing font size, but without spacing letters, had no impact on reading. In Experiment 2a, we addressed the role of a further potential confounding parameter, namely, eccentricity. Spacing makes some letters migrate to the periphery of the visual field and, thus, enter in a region of lower visual acuity. By moving non-spaced stimuli to lateral regions of the visual field, we pitted spacing and eccentricity against each other and concluded that the impact of spacing cannot be reduced to a spurious effect of peripheral vision. Experiment 2b replicated the results of Experiment 2a with a more natural reading task. In Experiment 3, we further separated spacing from word size by equating size across 4-, 6-, and 8-letter words. The results again showed that the effect of spacing cannot be reduced to an effect of physical size. Finally, in Experiment 4, we showed that the critical threshold of 2 spaces was constant across variations in font size.

Task and material

In the present experiments, our aim was to study the early, visual, component of word reading. To this end, we manipulated purely visual parameters (spacing, eccentricity, side) and measured their impact on lexical decision. We, therefore, expect that our conclusions should apply to any reading task sharing the same visual component, including more natural tasks such as reading aloud or reading for comprehension. Note that we used orthographically and phonologically plausible pseudowords in order to prevent any low-level response strategy. We have indications that our results do generalize beyond the lexical decision task. First, the main results of Experiment 2a were replicated in Experiment 2b using an overt reading task. Second, a spacing threshold of about 2 spaces was previously observed using a semantic decision task and also using overt reading in a patient with parietal damage (Vinckier et al., 2006).

In all experiments, we have used upper case stimuli. It might be argued that such format is relatively infrequent in daily life. However, there are converging indications that upper case words are not more difficult to read. Mean latencies do not differ between upper case and lower case words (Qiao et al., 2010). Moreover, the absence of a word length effect in normal reading conditions, as shown, e.g., in the present study, confirms that expert parallel reading prevails also with upper case words. Moreover, functional imaging studies have evidenced subliminal cross-case priming (e.g., radio > RADIO) in the ventral visual system, suggesting that the case quickly becomes irrelevant starting from early visual stages of word processing (Dehaene et al., 2004).

Reading spaced letters: Physiological mechanisms

The fast and parallel reading performance whose development culminates in literate adults is thought to result from a progressive tuning of the ventral visual system (Dehaene et al., 2010). This training, however, is restricted to the familiar reading format, namely, horizontally printed strings of contiguous letters in the central and right parafoveal portions of the visual field. In order to cope with degraded or unfamiliar displays, including words with spaced letters, readers resort to serial scanning of word fragments, explaining both the overall slowing and the positive correlation with the number of letters. There is functional imaging and neuropsychological evidence that this compensation process is based upon parietal attention-related mechanisms. Above a threshold of about 2 spaces between consecutive letters, concomitant with performance reduction, there is a sudden increase in BOLD signal in bilateral posterior intraparietal areas that do not belong to the typical reading network (Cohen et al., 2008). The causal role of parietal cortex was demonstrated in a patient with bilateral parietal lesions, whose reading performance dropped dramatically as soon as letter spacing passed the very same threshold of about 2 spaces. Thus, both fMRI and neuropsychology suggest that slow reading with a length effect reflects the deployment of attention-dependent spatial scanning strategies under parietal guidance, a process that is triggered when spacing exceeds the capacity of the ventral cortex for parallel reading and invariant word recognition. Applied to the present experiment, this conclusion implies that the invariance of the ventral visual system for letter spacing collapses suddenly above a critical threshold value of about 2 spaces.

The bigram coding hypothesis

A threshold value slightly below 2 spaces matches an explicit prediction of the LCD framework (Dehaene et al., 2005). According to this model, detectors of single letters, with a local receptive field, converge to create the slightly larger receptive fields of open bigram detectors sensitive to the spatial configuration of two letters. As mentioned in the Introduction section, based on the increase of receptive fields in the IT cortex by a factor of about 2.5 from one neural level to the next (Rolls, 2000), the LCD model proposes that blank spaces of 2 spaces should be sufficient to disrupt bigram detectors (Dehaene et al., 2005), precluding parallel encoding of letters into larger units. This value is, thus, a plausible though approximate estimator of the limits of the letter grouping ability of the ventral pathway.

One important consequence of the present research is to support the hypothesis that the fast recognition of combinations of letters plays a central role at some stage in the coding of written words, to such an extent that interfering with this representation drastically impedes the parallel analysis of letter strings. It should, however, be noted that, strictly speaking, the present experiments cannot determine the exact nature of this combinatorial that is disrupted by spacing. It could be pairs of letters (bigrams) but also perhaps a subset of these (e.g., only consonant bigrams; Perea, Acha, & Carreiras, 2009) or even larger units such as morphemes. Bigram coding has been proposed to play an important role in several recent models of orthographic processing (Grainger, Granier, Farioli, Van Assche, & van Heuven, 2006; Grainger & Whitney, 2004; Whitney, 2001) and is supported by several empirical findings. The number of shared bigrams can explain the amount of priming for subliminal words and their substrings (e.g., the fact that "grdn" primes "garden"; although see also Davis & Bowers, 2006; Grainger et al., 2006; Grainger & Holcomb, 2009; Humphreys, Evett, & Quinlan, 1990; Peressotti & Grainger, 1999; Schoonbaert & Grainger, 2004). Bigram frequency is a strong predictor of the activation of the visual word form area, a part of the ventral visual cortex that houses an orthographic representation of letter strings (Binder, Medler, Westbury, Liebenthal, & Buchanan, 2006; Vinckier et al., 2007). There is also support for the notion that the reading system quickly parses visual strings into subsequences corresponding to morphemes such as frequent prefixes and suffixes (Burani, Marcolini, De Luca, & Zoccolotti, 2008; Christianson, Johnson, & Rayner, 2005; Frost, Deutsch, Gilboa, Tannenbaum, & Marslen-Wilson, 2000), even if these are only "pseudomorphemes" semantically inappropriate in the current word context (Longtin, Segui, & Hallé, 2003; Rastle, Davis, & New, 2004). Clearly, further research will be needed to determine the exact level of orthographic coding that is disrupted by spacing.

Size invariance of the reading threshold

Finally, why is the spacing threshold of about 2 spaces invariant for changes in the size of letters, as shown in Experiment 4? This issue is not explicitly dealt with by the LCD model, but two observations may clarify this point. First, according to the LCD model, there is, at the earliest levels of the neural hierarchy, a tolerance for small variations in the position and size of visual features. The progressive increase in this tolerance up to detectors for letters and bigrams should contribute to an overall size invariance of word recognition processes. Second, readers have actually been exposed to letters of various sizes, and different neurons may well have become tuned to letters and bigrams of various dimensions. This is in agreement with monkey data showing that anterior IT neurons have receptor fields ranging from 3° to 26° and appropriate for the detection of objects of different angular sizes (Op De Beeck & Vogels, 2000). In this view, size invariance emerges progressively at increasingly higher stages of the visual hierarchy, responsible for letter detection and beyond. This is in agreement with the increase of invariance for size along the posterior to anterior axis of the lateral occipital cortex (Eger & Kell, 2008), a visual area involved in invariant object recognition and that is abutting and partially overlapping with the VWFA. Considering both early and late sources of size invariance, the main explanation for the invariance of the spacing threshold is simply that the receptive fields of bigram detectors are more than twice larger than the receptive fields of the letter detectors from which they receive their input.

Conclusion

Invariance is a fundamental requirement in reading—we must be capable of identifying words in spite of major changes in size, location, and spacing. The detection of specific combinations of letters such as bigrams may crucially contribute to the progressive construction of an invariant representation that preserves the identity of the letter string (Dehaene et al., 2005; Grainger & Whitney, 2004). The present experiments, however, identify a clear limitation of this architecture: word recognition is only invariant across small changes in letter spacing, while larger spaces severely disrupt reading and impose a switch to a radically distinct serial processing mode.

Acknowledgments

This research was funded by the Agence Nationale pour la Recherche (ANR, CORELEX project).

Commercial relationships: none.

Corresponding author: Laurent Cohen.
Email: laurent.cohen@psl.aphp.fr.
Address: AP-HP, Department of Neurology, Hôpital de la Salpêtrière, Paris, 75013, France.

Footnote

[1] In order to determine whether the interaction of spacing and length depended only on the trials in which words were most difficult to read (i.e., long words with widely spaced letters), we run the LME analysis of Experiments 1 and 4, removing the three extreme conditions, i.e., 8-letter words with a spacing of 3 or 4 spaces and 6-letter words with a spacing of 4 spaces. The results of this analysis were essentially the same as with the full set of data. Particularly, the optimal model remains unchanged, and the interaction of spacing and length remains significant, demonstrating that this effect cannot be reduced to the extreme conditions of Experiment 1.

References

Aghababian, V., & Nazir, T. A. (2000). Developing normal reading skills: Aspects of the visual processes underlying word recognition. *Journal of Experimental Child Psychology, 76,* 123–150.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59,* 390–412.

Binder, J. R., Medler, D. A., Westbury, C. F., Liebenthal, E., & Buchanan, L. (2006) Tuning of the human left fusiform gyrus to sublexical orthographic structure. *Neuroimage, 33,* 739–748.

Bub, D. N., & Lewine, J. (1988). Different modes of word recognition in the left and right visual fields. *Brain and Language, 33,* 161–188.

Burani, C., Marcolini, S., De Luca, M., & Zoccolotti, P. (2008). Morpheme-based reading aloud: Evidence from dyslexic and skilled Italian readers. *Cognition, 108,* 243–262.

Christianson, K., Johnson, R. L., & Rayner, K. (2005). Letter transpositions within and across morphemes. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31,* 1327–1339.

Cohen, L., & Dehaene, S. (2009). Ventral and dorsal contributions to word reading. In M. S. Gazzaniga (Ed.), *Cognitive neuroscience* (4th ed., pp. 789–804). The MIT Press.

Cohen, L., Dehaene, S., Naccache, L., Lehéricy, S., Dehaene-Lambertz, G., Hénaff, M. A., et al. (2000). The visual word form area: Spatial and temporal characterization of an initial stage of reading in normal subjects and posterior split-brain patients. *Brain, 123,* 291–307.

Cohen, L., Dehaene, S., Vinckier, F., Jobert, A., & Montavont, A. (2008). Reading normal and degraded words: Contribution of the dorsal and ventral visual pathways. *Neuroimage, 40,* 353–366.

Davis, C. J., & Bowers, J. S. (2006). Contrasting five different theories of letter position coding: Evidence from orthographic similarity effects. *Journal of Experimental Psychology: Human Perception & Performance, 32,* 535–557.

Dehaene, S., Cohen, L., Sigman, M., & Vinckier, F. (2005). The neural code for written words: A proposal. *Trends in Cognitive Sciences, 9,* 335–341.

Dehaene, S., Jobert, A., Naccache, L., Ciuciu, P., Poline, J.-B., Le Bihan, D., et al. (2004). Letter binding and invariant recognition of masked words: Behavioral and neuroimaging evidence. *Psychological Science, 15,* 307–313.

Dehaene, S., Pegado, F., Braga, L. W., Ventura, P., Filho, G. N., Jobert, A., et al. (2010). How learning to read changes the cortical networks for vision and language. *Science, 3,* 1359–1364.

Eger, E., & Kell, C. A. (2008). Kleinschmidt A. Graded size sensitivity of object-exemplar-evoked activity patterns within human LOC subregions. *Journal of Neurophysiology, 100,* 2038–2047.

Ellis, A. W. (2004). Length, formats, neighbours, hemispheres, and the processing of words presented laterally or at fixation. *Brain and Language, 88,* 355–366.

Ferrand, L., & New, B. (2003). Syllabic length effects in visual word recognition and naming. *Acta Psychologica, 113,* 167–183.

Frost, R., Deutsch, A., Gilboa, O., Tannenbaum, M., & Marslen-Wilson, W. (2000). Morphological priming: Dissociation of phonological, semantic, and morphological factors. *Memory and Cognition, 28,* 1277–1288.

Gaillard, R., Naccache, L., Pinel, P., Clemenceau, S., Volle, E., Hasboun, D., et al. (2006). Direct intracranial, fMRI and lesion evidence for the causal role of left inferotemporal cortex in reading. *Neuron, 50,* 191–204.

Grainger, J., Granier, J. P., Farioli, F., Van Assche, E., & van Heuven, W. J. (2006). Letter position information and printed word perception: The relative-position priming constraint. *Journal of Experimental Psychology: Human Perception and Performance, 32,* 865–884.

Grainger, J., & Holcomb, P. J. (2009). An ERP investigation of orthographic priming with relative-position and absolute-position primes. *Brain Research, 1270,* 45–53.

Grainger, J., & Whitney, C. (2004). Does the huamn mnid raed wrods as a wlohe? *Trends in Cognitive Sciences, 8,* 58–59.

Hasson, U., Levy, I., Behrmann, M., Hendler, T., & Malach, R. (2002). Eccentricity bias as an organizing principle for human high-order object areas. *Neuron, 34,* 479–490.

Humphreys, G. W., Evett, L. J., & Quinlan, P. T. (1990). Orthographic processing in visual word identification. *Cognitive Psychology, 22,* 517–560.

Lavidor, M. (2002). An examination of the lateralized abstractive/form specific model using MiXeD-CaSe primes. *Brain and Cognition, 48,* 413–417.

Lavidor, M., & Ellis, A. W. (2002). Word length and orthographic neighborhood size effects in the left and right cerebral hemispheres. *Brain and Language, 80,* 45–62.

Legge, G. E., Ahn, S. J., Klitz, T. S., & Luebker, A. (1997). Psychophysics of reading: XVI. The visual span in normal and low vision. *Vision Research, 37,* 1999–2010.

Longtin, C.-M., Segui, J., & Hallé, P. A. (2003). Morphological priming without morphological relationship. *Language and Cognitive Processes, 18,* 313–334.

Nazir, T. A., Jacobs, A. M., & O'Regan, J. K. (1998). Letter legibility and visual word recognition. *Memory and Cognition, 26,* 810–821.

New, B., Pallier, C., Brysbaert, M., & Ferrand, L. (2004). Lexique 2: A new French lexical database. *Behavioral Research Methods, Instruments, and Computer, 36,* 516–524.

Op De Beeck, H., & Vogels, R. (2000). Spatial sensitivity of macaque inferior temporal neurons. *Journal of Computerized Neurology, 426,* 505–518.

O'Regan, J. K., Levy-Schoen, A., Pynte, J., & Brugaillere, B. (1984). Convenient fixation location within isolated words of different length and structure. *Journal of Experimental Psychology: Human Perception and Performance, 10,* 250–257.

Perea, M., Acha, J., & Carreiras, M. (2009). Eye movements when reading text messaging. *Quarterly Journal of Experimental Psychology, 62,* 1560–1567.

Peressotti, F., & Grainger, J. (1999). The role of letter identity and letter position in orthographic priming. *Perception & Psychophysics, 61,* 691–706.

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, 15,* 1786–1799.

Rastle, K., Davis, M. H., & New, B. (2004). The broth in my brother's brothel: Morpho-orthographic segmentation in visual word recognition. *Psychonomic Bulletin and Review, 11,* 1090–1098.

Rayner, K., & Bertera, J. H. (1979). Reading without a fovea. *Science, 206,* 468–469.

Rolls, E. T. (2000). Functions of the primate temporal lobe cortical visual areas in invariant visual object and face recognition. *Neuron, 27,* 205–218.

Schoonbaert, S., & Grainger, J. (2004). Letter position coding in printed word perception: Effects of repeated and transposed letters. *Language and Cognitive Processes, 19,* 333–367.

Serre, T., Oliva, A., & Poggio, T. (2007). A feedforward architecture accounts for rapid categorization. *Proceedings of the National Academy of Sciences of the United States of America, 104,* 6424–6429.

Vinckier, F., Dehaene, S., Jobert, A., Dubus, J. P., Sigman, M., & Cohen, L. (2007). Hierarchical coding of letter strings in the ventral stream: Dissecting the inner organization of the visual word-form system. *Neuron, 55,* 143–156.

Vinckier, F., Naccache, L., Papeix, C., Forget, J., Hahn-Barma, V., Dehaene, S., et al. (2006). "What" and "Where" in word reading: Ventral coding of written words revealed by parietal atrophy. *Journal of Cognitive Neuroscience, 18,* 1998–2012.

Weekes, B. S. (1997). Differential effects of number of letters on word and nonword naming latency. *Quarterly Journal of Experimental Psychology, 50A,* 439–456.

Whitney, C. (2001). How the brain encodes the order of letters in a printed word: The SERIOL model and selective literature review. *Psychonomic Bulletin and Review, 8,* 221–243.

Whitney, C., & Lavidor, M. (2004). Why word length only matters in the left visual field. *Neuropsychologia, 42,* 1680–1688.