The influence of semantic context on initial eye landing sites in words

Frédéric Lavigne a,*, Françoise Vitu b, Géry d’Ydewalle c

a Laboratoire de Psychologie Expérimentale et Quantitative, Université de Nice – Sophia Antipolis,
24 avenues des Diables bleus, 06357 Nice Cedex 4, France

b CNRS, Université René Descartes, Paris, France

c University of Leuven, Leuven, Belgium

Received 17 December 1998; received in revised form 5 January 2000; accepted 6 January 2000

Abstract

To determine the role of ongoing processing on eye guidance in reading, two studies examined the effects of semantic context on the eyes’ initial landing position in words of different levels of processing difficulty. Results from both studies clearly indicate a shift of the initial fixation location towards the end of the words for words that can be predicted from a prior semantic context. However, shifts occur only in high-frequency words and with prior fixations close to the beginning of the target word. These results suggest that ongoing perceptual and linguistic processes can affect the decision of where to send the eyes next in reading. They are explained in terms of the easiness of processing associated with the target words when located in parafoveal vision. It is concluded that two critical factors might help observing effects of linguistic variables on initial landing sites, namely, the frequency of the target word and the position where the eyes are launched from as regards to the beginning of the target word. Results also provide evidence for an early locus of semantic context effects in reading. © 2000 Elsevier Science B.V. All rights reserved.

MSC: 2300; 2340; 2343

Keywords: Context effects; Eye movements; Priming; Semantic priming; Sentence reading

* Corresponding author.
E-mail address: lavigne@unice.fr (F. Lavigne).
1. Introduction

In reading, the eyes progress on the line of text with forward saccades of variable sizes. Two out of three words are fixated with a forward saccade, some words being skipped during a first eye pass. The position where the eyes initially land in a word is generally located between the beginning and the middle of the word, but in some instances it can be located at the beginning or end of the word (Rayner, 1979).

During the last 30 years, a large amount of research has been devoted to understanding what determines the variability in both word skipping rate and initial landing sites in words. Results as a rule show that visuomotor factors associated with saccadic programming (such as the length of the next word and the distance from the word where the eyes are launched from) contribute largely to this variability (O’Regan, 1979; McConkie, Kerr, Reddix & Zola, 1988; Vitu, O’Regan & Mittau, 1990; Vitu, 1991a,b; Vitu, O’Regan, Inhoff & Topolski, 1995; Radach & McConkie, 1998). In contrast, linguistic factors such as the processing difficulty of the words located in parafoveal vision seem to participate to a smaller extent in the variability. As noted by Brysbaert and Vitu (1998) (O’Regan, 1990), most studies have failed to find an influence of ease of processing a parafoveal word on the likelihood of skipping it, and when effects are reported, they are either small in amplitude or can be attributed to some other variables.

Concerning initial landing sites in words, Beauvillain, Doré and Baudoin (1996) reported a difference in initial landing sites of less than a letter between words that differ in the frequency of their initial letter sequence. On the other hand, Everatt and Underwood (1992) showed an initial landing difference as a function of where the uniquely identifying information in the word is located. But this effect could never be repeated, and it has been attributed to either artifacts of the eye tracking system (Rayner & Morris, 1992) or confounding variables such as the luminance of the letters (Baccino & Pynte, 1994). Finally, Dubois and Sprenger-Charolles (1988) reported an effect of semantic context on initial landing sites, when paired comparisons were made between sentences that were identical except for the target word (that was or was not semantically related to a preceding prime word). Since target words differed in both the visual attributes and frequency of occurrence in the language, conclusions can hardly be drawn as to the role of semantic factors.

At the same time, there is evidence in the literature that processing of parafoveal words occurs at least up to a lexical level (Rayner & Pollatsek, 1989; Rayner & Balota, 1989; Rayner & Morisson, 1981). Both naming time and gaze duration (or total time the eyes spend on the word) are shorter when the word was visible in parafoveal vision before being fixated than when it was masked. In addition, the resulting parafoveal preview benefit is larger for high- than low-frequency words (Inhoff & Rayner, 1986; Vitu, 1991b) and for words that are predictable from the prior semantic and/or syntactic context (Balota, Pollatsek & Rayner, 1985; McClelland & O’Regan, 1981; Paap & Newsome, 1981; Vitu, 1991b). Whether there is a semantic preprocessing of the parafoveal words is still an open question (Rayner, 1979; Rayner, McConkie & Zola, 1980; Balota & Rayner, 1983, 1991; Kennison & Clifton, 1995; Lavigne & Dubois, 1999).
Thus, although some information is extracted from words located in parafoveal vision, it does not seem to have a clear effect on where to move the eyes next. According to O’Regan (1990) (see also O’Regan & Lévy-Schoen, 1987), the paradox results from both the slowness of parafoveal word identification processes and the use of an autonomous scanning strategy that is independent of these processes and relies only on low-level visual processes. The eyes could be guided through the line of text by a strategy attempting to drive the eyes from the center of one word to the center of the next word. Variability in both word skipping rate and initial landing sites then result mostly from low-level visuomotor factors that affect saccadic accuracy; the rare instances in which ongoing linguistic processes influence the eye landing could correspond to cases where the prior fixation duration is longer than usual.

This oculomotor-control view is challenged by a more cognitive view, that suggests that parafoveal word processing and the programming of the next saccade overlap in time, and that ongoing linguistic processes can modify the planned saccade (Morrison, 1984; Henderson & Ferreira, 1990; see also Reichle, Pollatsek, Fisher & Rayner, 1998; Rayner, Reichle & Pollatsek, 1998). According to this view, all inter-word saccades initially aim for the center of the next word, but they deviate from the target when processing of the word is terminated before the saccade is computed. Thus, the likelihood that a saccade is influenced by ongoing processes then depends on the ease of processing associated with the parafoveal word.

The purpose of the present study was to distinguish between oculomotor- and cognitive-control views of eye movements in reading, and to determine the extent to which initial landing sites in words are sensitive to ongoing perceptual and linguistic processes. In particular, the present experiments tested whether the position where the eyes initially land in a word depends on the word’s predictability from a prior semantic context and whether the likelihood of observing such an effect is a function of the ease of processing associated with the parafoveal word. If context effects are obtained in conditions that facilitate parafoveal word processing, then this argues against the hypothesis that inter-word saccades are determined by an autonomous scanning strategy that is independent of ongoing linguistic processes (O’Regan, 1990). At the same time, this would provide further evidence for the case of semantic parafoveal word processing (Fuentes & Tudela, 1992; Fuentes & Ortelles, 1993; Fuentes, Carmona, Agis & Catena, 1994; Lavigne & Dubois, 1999; see also McClelland & O’Regan, 1981), and it would bring some insights on the question of where in the time course of word identification semantic processes intervene. This question has indeed been extensively debated in the last 20 years, and it is still not clear whether semantic context effects emerge before or after lexical access has been completed (Swinney, 1979; Onifer & Swinney, 1981; Fodor, 1983; Marslen-Wilson & Tyler; 1987; Tabossi, 1988a,b; Pacht & Rayner, 1993; Tabossi & Zardon, 1993). As a saccade is programmed within 250 ms of an average fixation duration in reading, linguistic effects on saccades then suggest that semantic context effects can be shown at an earlier stage than in lexical decision and naming tasks, which take about 500 ms and which, in addition, involve post-access strategies (Forster, 1981; Balota & Chumbley, 1984;
In the present experiments, eye movement data were recorded in two experiments that were part of a larger study (Lavigne, Vitu & d'Ydewalle, submitted). Participants were presented isolated sentences with one or two semantically (un)related prime(s) preceding a target word. In both studies, conditions were selected to maximize the occurrence of early context effects (i.e., on initial landing sites). First, we used for related prime(s) and target words, pairs of words that were strongly associated. Second, we used both high- (Experiments 1 and 2) and low-frequency target words (Experiment 1 only). It has been shown in prior studies that high-frequency words are more likely to be preprocessed in parafoveal vision (Inhoff & Rayner, 1986; Vitu, 1991b), and therefore they may enhance the occurrence or the amplitude of context effects (Shubert & Eimas, 1977; Becker, 1979, 1985; Vitu, 1991b). Third, we determined a posteriori the launch site (i.e., the last position of the eyes before the target word), since context effects may emerge only for close launch sites that favor parafoveal preprocessing (McClelland & O'Regan, 1981; Vitu, 1991a,b). In addition, to ensure that context effects could not be the result of the target words having different visual attributes, pairs of, respectively, related and unrelated sentences contained the same target word, but differed by the preceding prime being used.

2. Experiment 1

2.1. Method

Participants. Eighteen students (11 from the University of Leuven, Belgium, and seven from Paris, France) who were between 18 and 30 years old, were paid to participate in the experiment at the University of Leuven, Belgium. They were all native French speakers (first language) and had normal or corrected-to-normal vision (only with glasses).

Material. Two sets of 72 sentences were constructed. Each sentence contained a prime and a target word. In half the sentences, both words were semantically related, and in the other half, both words were unrelated. Both related and unrelated sentences were matched except for the prime word which was either semantically related to the target or not. The related and unrelated primes were matched in length up to a two-letter difference.

The choice for related prime and targets was made following results from a preliminary study in which 20 participants were presented with a list of 150 words and asked to give for each of them, all words which appeared to them as being semantically associated to the word (following the order the words as it came to their mind). Seventy-two word pairs were selected according to the participants’ responses: for a given word (the prime), a same word (the target word) had to be cited by at least 75% of the participants. In unrelated pairs, the prime, though fitting into the sentence, did not have any semantic relation with the target.
All selected target words were between 6 and 8 letters long. Half of them were of high frequency of occurrence in the language (more than 53 utterances per million, with a mean of 92 occurrences per million) and the other half of low frequency (less than 16 utterances per million, with a mean of 9 occurrences per million; Trésor de la Langue Française, 1971). Target words did not have strong competitive meanings, which are known to influence word processing time (Jastrzembski & Stanners, 1975; Jastrzembski, 1981; Gernsbacher, 1984) and to lead to pseudo-frequency effects (Millis & Button, 1989; Masson, 1991, 1995). In order to prevent neighborhood frequency effects (Grainger, 1988, 1990), none of them had orthographical neighbors.

In all experimental sentences, the prime and the target corresponded, respectively, to the subject and object nouns of the sentence and were included in a single clause sentence. The target was never the final word in the sentence and was always followed by an adjectival noun phrase, and it was never preceded or followed by a punctuation mark. In half of the sentences, the ending noun phrase contained a coherent word (as regards to the rest of the sentence) and the other half contained an incoherent word. The ending coherent or incoherent word could not be seen while fixating the target word (since it was located on average 15 characters to the right of the end of the target word) and could not therefore interfere with the reading of the previous part of the sentence.

The two conditions of semantic relatedness for a given target word (a high-frequency target: trousers; or a low-frequency target) were as indicated in the following examples. (The original sentences were in French. The prime and the target words are italicized.)

1. La ceinture tenait le pantalon du vieillard. (The belt held the trousers of the old man: related context, high-frequency target.)
2. La ficelle tenait le pantalon du vieillard. (The string held the trousers of the old man: unrelated context, high-frequency target.)
3. La momie provenait d’une pyramide d’Egypte. (The mummy was from a pyramid of Egypt: related context, low-frequency target.)
4. Le miroir provenait d’une pyramide d’Egypte. (The mirror was from a pyramid of Egypt: unrelated context, low-frequency target.)

The level of semantic relatedness within the sentences between the prime and the target was tested for each sentence in a preliminary experiment where 15 participants were presented with the beginning of each experimental sentence (up to the target word). The participants’ task was to indicate which word came first to their mind when reading the initial parts of the sentences. Results showed on average that the associated target word was cited in 94% of the cases in the related condition, and only in 3% of the cases in the unrelated condition.

**Design.** There were two levels of semantic relatedness between the prime and the target (related vs. unrelated) and two levels of target frequency (high vs. low frequency). Each participant saw all experimental conditions and all target words; over all participants, all target words were seen in all conditions, Latin Square design being used. Thus, each participant saw only 72 experimental sentences (half being related and the other half unrelated). The 72 experimental sentences were
mixed with 240 filler sentences which were similar to those used by Vitu (1991b) and were used for another reading experiment, yielding a total of 312 sentences. All sentences were presented in a random order for each participant within two blocks of trials, each containing the same number of experimental sentences (156 sentences).

**Apparatus.** Sentences were displayed on a VGA monitor screen (refreshed every 14 ms) in lower cases (except for the first letter of each sentence). Each letter subtended a third of a degree of visual angle. Eye movements were monitored via a 5.5 generation Dual Purkinje eyetracker, connected to two PC compatible micro-computers. The first computer presented sentences and recorded the participant’s response for each sentence, and the second sampled the successive eye positions every millisecond. Eye movements were monitored from the right eye, although viewing was binocular. The system accuracy corresponded to about an eighth of a character. Luminance on the screen was adjusted to a comfortable level throughout the experiment. The room was dark, except for a dim indirect light source.

**Procedure.** Upon arrival, each participant was seated comfortably at a distance of 60 cm from the computer screen. His or her head movements were minimized with a bite-bar and a head and chin rest. The calibration phase then took place which required the participant to fixate as accurately as possible (while pressing a key button) a fixation point which appeared at 15 successive positions distributed over the screen along the two diagonals and the middle horizontal line (where the sentence was to be further displayed). As soon as the participant pressed a key button, the point disappeared and reappeared at the next screen position. At the moment the participant pressed the key button, the eye position was recorded. The calibration phase was repeated until the correlation between the different positions of the points on the screen and the corresponding eyes’ locations was better than 0.95. After calibration, participants were given 12 practice trials followed by a total of two blocks of 156 test trials (participants were allowed to pause whenever they wanted). Each session lasted approximately 45 min.

At the beginning of each trial, the participant was asked to fixate a small vertical segment located on the left part of the screen. At that moment, a dot indicating the eyes’ position and moving in synchrony with the eyes was visible on the screen. When a fixation was detected in a region of plus or minus half a character around the vertical segment, a sentence appeared on the screen with its third character located on the vertical segment’s position. The sentence remained on the screen until the participant pressed a button to indicate that he or she was ready with reading of the sentence. The sentence then disappeared and two boxes appeared at the bottom of the screen, indicating the two possible responses (coherent or incoherent sentence). The participant was asked to fixate one of the two boxes in order to indicate his or her response. When the fixated box was highlighted, the participant pressed the button again to confirm his or her response. Then the next trial started.

**Data selection.** In all analyses, data were selected when there was no blink or any signal irregularity before, or during the reading of the target word. Furthermore, to prevent the reading of the end of the sentence interfering with the initial reading of the target word, only data corresponding to the first eye pass on the word and
preceded by a progressive saccade were considered for the analysis. The proportion of errors to the coherence question of the sentence was small (less than 9% of the trials). Since participants reported that ‘incoherent’ sentences could be interpreted as coherent (such as in the sentence finishing by ‘a chocolate tea pot’, for example), incorrectly responded sentences were not excluded from the analysis.

The eyes’ initial landing position in the target word was measured as a function of the prime–target semantic relatedness and the target word frequency. The launch site (i.e., the last eye position before fixating the target word) was defined a posteriori and was introduced as a third factor in the analyses of variance. The values in all analyses of variance corresponded to the mean obtained for each participant in each condition, so that each participant’s contribution to the global mean was equivalent. The distributions of initial landing sites obtained in the different conditions are presented for each analysis made, but no statistical analysis was made directly on these distributions.

3. Results and discussion

Global and local eye-movement characteristics. For comparison with previous eye movement studies, we first analyzed the eye movement pattern globally. Results show that progressive and regressive saccades were about 10 and 5 letters, respectively, with 17% regressive saccades. Average fixation durations and gaze durations were, respectively, of 281 and 347 ms. Sixty-seven percent of the words were fixated and 21% were refixated.

Local analyses revealed that saccades directed to the target words were longer on average (12.4 letters) than saccades to other words in the sentences, a fact which probably resulted from the fact that target words were quite long (from 6 to 8 letters; O’Regan, 1979). Only 4% target words were skipped.

Furthermore, gaze durations on the target words were shorter for related (329 ms) compared to unrelated targets (353 ms) and for high- (324 ms) compared to low-frequency target words (358 ms). Single fixation durations were also shorter for related (288 ms) than unrelated targets (322 ms) and for high- (287 ms) than low-frequency targets (310 ms).

Initial landing site in target words. Figs. 1(a) and (b) present the distribution of initial landing sites in target words for both conditions of semantic relatedness and both target frequencies. The figures show that the eyes mostly land between the beginning and the middle of the words in all four conditions, and that the distributions for related and unrelated targets largely overlap. The average landing site was very similar for related and unrelated targets, both for high- (letter positions 2.61 vs. 2.53, respectively) and for low-frequency target words (2.68 vs. 2.73, respectively). There were no significant effects of semantic relatedness and frequency, and no significant interaction.

As a result of visual acuity constraints, the amount of information that can be extracted from a parafoveal word is a function of how far the eyes are located from the beginning of the word: the further the eyes from the beginning of the word, the
less likely the word can be identified (Brysbaert, Vitu & Schroyens, 1996). According to measurements of the perceptual span, no more than 6–8 letters to the right of the fixation point can be used for semantic processing (McConkie & Rayner, 1975; see also Rayner, 1975a,b; Rayner, Well & Pollatsek, 1980; Rayner & Pollatsek, 1981; Rayner, Well, Pollatsek & Bertera, 1982). Since in the present study saccades to the target word were 12.4 letters long on average, the eyes were perhaps on most occasions launched from too far from the beginning of the word to allow parafoveal preprocessing to occur and context effects to emerge (McClelland & O’Regan, 1981). Further analyses indeed revealed that on average the eyes were launched from a distance of 9.8 letters from the beginning of the target word, with almost half the launch sites being larger than 10 letters, therefore suggesting that in at least 50% of the cases, no information could be extracted from the parafoveal word. However, if initial landing sites in words are sensitive to ongoing word processing in parafoveal vision, then an effect of semantic factors should emerge at least in cases where the launch site is below the limits of the perceptual span.

We tested this hypothesis by introducing the launch site as a third factor in the analysis. Two classes of launch sites were defined a posteriori that contained almost equal numbers of data in all conditions and that supposedly provided different amounts of parafoveal word information. Close launch sites included launch sites up
to eight characters from the space preceding the target word, and far launch sites corresponded to the remaining cases.

Figs. 2(a)–(d) present the distribution of initial landing sites in related vs. unrelated target words, separately for high- and low-frequency targets, and for close and far launch sites. First, the figures show the classical launch site effect originally obtained by McConkie et al. (1988) (see also Vitu et al., 1995): initial landing sites are shifted towards the end of the words for close compared to far launch sites. Second, there is a large overlap of the distributions between related and unrelated cases but only for far launch sites. For close launch sites, initial landing sites in high-frequency words are shifted towards the end of the words in the related condition compared to the unrelated condition; initial landing sites in low-frequency targets show the reverse pattern.

Analyses of variance showed a significant effect of launch site (close vs. far) \(F(1, 17) = 27.03, \text{MSE} = 1.26, p < 0.001\), but no significant effect of relatedness or frequency \(F < 1\). Two-way interactions were not significant. In contrast, the three-way interaction between relatedness, frequency, and launch site was significant \(F(1, 17) = 5.23, \text{MSE} = 0.56, p < 0.05\). For high-frequency words, the effects of launch site, relatedness, and the interaction between launch site and relatedness were significant \(F(1, 17) = 16.43, \text{MSE} = 1.09, p < 0.001; F(1, 17) = 4.48, \text{MSE} = 0.37, p < 0.05, \) and \(F(1, 17) = 8.15, \text{MSE} = 0.44, p < 0.01, \) respectively). For low-frequency words, the only factor that showed a significant effect was the launch site \(F(1, 17) = 21.83, \text{MSE} = 0.74, p < 0.001\).

---

**Fig. 2.** Distributions of the initial eye landing positions in the target words (in letters) as a function of the semantic relatedness between the target and a previously read prime (related vs. unrelated), for close ((a), (c)) and far launch sites ((b), (d)) (or, respectively, up to 9 characters from the beginning of the target word, and above 9 characters from the beginning of the target word), and high- ((c), (d)) vs. low-frequency words ((a), (b)).
Thus, the eyes are sent further away into words that can be predicted from the prior semantic context, but only in cases where the eyes’ launch site is relatively close to the word and the word is of high frequency. However, since target words were of variable lengths (between 6 and 8 letters), there is still a possibility that this effect results from different word lengths in the different conditions, due to the post-hoc partitioning of the data. To check for this possibility, initial landing sites in words were re-coded with the following formula that takes word length into account: $f = (f - 0.5)/wl$, where $f$ is the fixation location, and $wl$ is the word length (Vitu et al., 1995). Then they were divided into three intervals of equal sizes (1/3). The distributions of re-coded landing sites confirm the conclusions from the letter landing site analysis. Results in Figs. 3(a)–(d) show a clear shift of the initial landing sites towards the end of the words in related cases compared to unrelated cases, but only for high-frequency targets that were launched from a close distance (see Fig. 3(c)).

The analysis of variance shows a significant effect of launch site ($F(1,17) = 31.38$, $MSE = 0.14, p < 0.001$), and a significant three-way interaction between launch site, relatedness, and frequency ($F(1,17) = 6.92$, $MSE = 0.13, p < 0.001$). Specific comparisons show that for high-frequency targets the effects of launch site, relatedness, and the interaction between launch site and relatedness are significant ($F(1,17) = 26.23$, $MSE = 0.11, p < 0.0005$; $F(1,17) = 5.09$, $MSE = 0.09, p < 0.05$, and $F(1,17) = 8.41$, $MSE = 0.11, p < 0.01$, respectively).

Fig. 3. Distributions of the initial eye landing positions in the target words as a function of the semantic relatedness between the target and a previously read prime (related vs. unrelated), for close ((a), (c)) and far launch sites ((b), (d)), and high- ((c), (d)) vs. low-frequency words ((a), (b)). Close launch sites corresponded to cases in which the last fixation before the initial fixation on the target word was located 8 or less characters from the space in front of the beginning of the target word, and far launch sites corresponded to the remaining cases (above 8 characters). Initial fixation positions in the target words have been re-coded using the following formula that takes word length into account: $f = (f - 0.5)/wl$, where $f$ is the fixation location and $wl$ is the word length. Re-coded initial landing sites were then divided into three zones of equal sizes (1/3) that are called, respectively, beginning (BEG), middle (MID), and end (END).
Prior fixation durations. According to the oculomotor-control view, ongoing processes associated with parafoveal word identification cannot intervene with saccadic programming and therefore cannot influence initial landing sites, unless the prior fixation duration is longer than usual (O’Regan, 1990). Thus, it might be that the effects of semantic context observed for high-frequency targets that were launched from a close distance resulted from longer prior fixation durations in that condition.

Table 1 shows the mean durations of the fixation that preceded the initial fixation on the target word for related and unrelated cases, for high- and low-frequency target words, and for close and far launch sites. The results indicate, contrary to the prediction made by the oculomotor-control view, that prior fixation durations tended to be shorter and not longer in the high-frequency and related conditions. Furthermore, fixation durations were shorter for close compared to far launch sites. The analysis of variance on these fixation durations shows a significant effect of launch site ($F(1, 17) = 7.31$, $MSE = 3576$, $p < 0.05$), and a marginally significant effect of relatedness ($F(1, 17) = 3.32$, $MSE = 1750$, $p < 0.10$), with still a significant effect of relatedness in the case of high-frequency words ($F(1, 17) = 5.17$, $MSE = 1635$, $p < 0.05$). None of the two- and three-way interactions were significant.

As can be inferred from Table 1, in conditions that favor an effect of semantic context, namely, when the target word is related and of high frequency and when the eye is launched from a close distance, prior fixation durations tend to be shorter and not longer than in the other conditions. This pattern of results clearly contradicts the predictions made by the strategy–tactics theory (O’Regan, 1990). The tendency for a decrease of fixation durations with semantic relatedness (at least for high-frequency words and close launch sites) actually suggests that semantic context does not only affect the initial eye landing position in the target word or the time spent fixating the target, but also the time spent on the prior fixation. This favors the idea of a parallel processing of both foveal and parafoveal word information (Kennedy, 1998; Schroyens, Vitu, Brysbaert & d’Ydewalle, 1999; see also for opposite findings, Henderson & Ferreira, 1993; Rayner, Fischer & Pollatsek, 1998). However, further evidence will be needed to conclude in favor or against a parallel vs. serial hypothesis for the processing of both foveal and parafoveal words.

Table 1
Mean prior fixation duration (i.e., the last fixation that preceded the initial fixation on the target word) as a function of both semantic relatedness between the target and a previously read prime (related vs. unrelated), and target word frequency (low vs. high), for close and far launch sites (Experiment 1)

<table>
<thead>
<tr>
<th>Number of related primes</th>
<th>Launch site</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Close</td>
<td>Far</td>
<td></td>
</tr>
<tr>
<td>HF-related</td>
<td>215</td>
<td>255</td>
<td></td>
</tr>
<tr>
<td>HF-unrelated</td>
<td>253</td>
<td>250</td>
<td></td>
</tr>
<tr>
<td>LF-related</td>
<td>235</td>
<td>278</td>
<td></td>
</tr>
<tr>
<td>LF-unrelated</td>
<td>240</td>
<td>268</td>
<td></td>
</tr>
</tbody>
</table>
Conclusion. The results show that semantic context can influence the position where the eyes initially land in a word. This occurs when parafoveal word processing is facilitated, or when the eyes are close enough to the beginning of the word on a prior eye fixation, and the word is of high frequency. This result, which contrasts with previous findings (O’Regan, 1990; Brysbaert & Vitu, 1998), suggests that there are at least two critical factors that constrain semantic context effects. First, the location of the eyes on a prior eye fixation determines how much information from the word can be extracted in parafoveal vision. Second, the frequency of the target word affects the speed with which the word is lexically accessed. The same conclusion was reached with studies that measured the joint effects of semantic context, parafoveal preview, and word frequency on naming times or gaze durations (Balota et al., 1985; Vitu, 1991b). The level of semantic association between the prime and the target is perhaps another critical factor for the occurrence of context effects. Indeed, one characteristic of the present study in comparison with previous research was the use of strongly associated word pairs.

The finding that semantic processing can occur at early stages of word processing, namely, while the word is still in parafoveal vision (Everatt & Underwood, 1992; Fuentes & Tudela, 1992; Fuentes & Ortells, 1993; Fuentes et al., 1994; Lavigne & Dubois, 1999; see also McClelland & O’Regan, 1981), suggests that anticipations from the previous context influence word processing from its very first moments while only partial visual information as to the letters that constitute the word is available. This interpretation is compatible with a context-dependent view (Tabossi, 1988a,b; Tabossi & Zardon, 1993) and with hypotheses of anticipatory forward context effects such as automatic spreading activation and expectancy mechanisms (Posner & Snyder, 1975a,b; Anderson, 1983; Neely & Keefe, 1989; Neely, 1991).

4. Experiment 2

In Experiment 1, semantic context was manipulated by using pairs of strongly associated prime and target words. Experiment 2 was designed to replicate results of Experiment 1 with a different procedure. Here, a given target word was preceded by zero, one or two (un)related primes. If word predictability is a critical factor in the occurrence of context effects on initial landing sites, then the findings of Experiment 1 should be replicated or even appear in a clearer manner because context effects are larger when several primes precede the target word (Brodeur & Lupker, 1994; Balota & Paul, 1996; Lavigne & Vitu, 1997). Since Experiment 1 has shown the presence of semantic context effects only for high-frequency words, all target words in Experiment 2 were of high frequency.

4.1. Method

Participants. Eighteen students (12 from the University of Leuven, Belgium, and six from Paris, France), between 18 and 30 years old, were paid to participate in the
experiment which was carried out at the University of Leuven, Belgium. They were all native French speakers and had normal or corrected-to-normal vision (only with glasses).

**Materials.** Three sets of 72 sentences were constructed. Each contained a prime, an interposed word, and a target word. In the first set of sentences, the prime and target were semantically related, and the interposed word was semantically related with the prime and target (2-prime condition). In the second set, the prime and target were related, but the interposed word was not related with the prime and target (1-prime condition). In the third set, the prime, interposed word, and target were unrelated (0-prime condition). The 72 0-prime sentences were matched with the 72 corresponding 1-prime and the 72 corresponding 2-prime sentences, being exactly identical, except for the prime and interposed words. The related and unrelated primes and interposed word were matched in length up to a 2-letter difference.

The choice for related prime, interposed word, and targets was based on the same preliminary study as for Experiment 1. After analysis of participants’ productions in the preliminary experiment, 72 prime–target pairs and 72 interposed word–target pairs were selected. The pairs corresponded to cases in which the participants’ responses were most coherent: for a given word (the prime), the same word (the target word) had been cited by at least 70% of the participants. In cases where the prime or the interposed word was unrelated to the target, there was no semantic relation with the target, but still fitting into the sentence.

All selected target words were between 5 and 7 letters long and they were all of high frequency of occurrence in the language (more than 44 utterances per million, with a mean of 83 occurrences per million; Trésor de la Langue Française, 1971). They did not have strong competitive meanings. Furthermore, none of them had orthographic similar neighbors. In all experimental sentences, the prime, interposed word and target were included in a single clause sentence. The target was never the final word in the sentence and was always followed by an adjectival noun phrase. It was also never preceded or followed by a punctuation mark. The 72 experimental sentences were mixed with 72 filler sentences. This was made to ensure that the regular syntactic structure of the experimental sentences would not lead participants to adopt specific reading strategies. In the filler sentences, the ending noun phrase contained a coherent word (as regards to the rest of the sentence) and the other half contained an incoherent word. The ending coherent or incoherent word could not be seen while fixating the target word (since it was located on average 15 characters to the right of the end of the target word) and could not therefore interfere with the previously read part of the sentence.

The three conditions of semantic relatedness for a given target word were as indicated in the following examples. (The original sentences were in French. The prime, interposed, and target words are italicized.)

1. *Le couloir et les fenêtres amélioraient agréablement la petite maison de briques.*
   (The *corridor* and the *windows* nicely brightened up the little brick *house*: 2-prime context.)
2. Le couloir et les fleurs égayaient agréablement la petite maison de briques. (The corridor and the flowers nicely brightened up the little brick house: 1-prime context.)

3. Les büches et les fleurs égayaient agréablement la petite maison de briques. (The lumber and the flowers nicely brightened up the little brick house: 0-prime context)

The level of semantic relatedness within the sentences between the prime–interposed word and the target was tested for each sentence in a second preliminary experiment. Fifteen students from Paris who did not run the first preliminary experiment were presented with the beginning of each experimental sentence (up to the target word). The participants’ task was to indicate which word came first to their mind when reading the initial parts of the sentences. The associated target word was cited in 86% of the cases in the 2-prime condition, 44% in the 1-prime condition, and 7% of the cases in the 0-prime condition.

**Design.** There were three levels of semantic relatedness between the prime and the target (2-prime vs. 1-prime vs. 0-prime). Each participant saw all experimental conditions and all target words (Latin Square design). Each participant saw a total of 144 sentences, with 72 experimental sentences (24 in each condition) and 72 filler sentences. All sentences were presented in a random order for each participant, within a single block of trials.

**Apparatus and procedure.** The same apparatus and procedure were used as in Experiment 1.

**Data selection.** The same criteria as in Experiment 1 were used. The proportion of errors to the question of the coherency of the sentence was small (less than 14% of the trials). Thus, good and false response trials were again considered in the analyses.

The eyes’ initial landing position in the target word was measured as a function of semantic relatedness. The analysis of variance was carried out on the mean obtained for each participant in each condition.

**5. Results and discussion**

**Global and local eye-movement characteristics.** Participants made 13% regressive saccades; the average size of both progressive and regressive saccades was about 9 and 7 letters, respectively. Fixation and gaze durations were on average 297 and 346 ms, respectively. Seventy-one percent of the words were fixated and 15% were re-fixated. The size of the saccades to the target words was 9.2 letters on average and only 3% of the target words were skipped.

Gaze durations on the target words were longer in the 0-prime condition (389 ms) than in both 1- (364 ms) and 2-prime conditions (359 ms). The same was true for single fixation durations (368, 339, and 336 ms, respectively).

**Landing site in target words.** In a first analysis, we measured the mean initial landing sites in the target words for the three conditions of semantic relatedness. Results show that the eyes land between the beginning and the middle of the words in all conditions of semantic relatedness, with the eyes landing only slightly further
away in 1- and 2-prime conditions (2.7 and 2.8 characters) than in the 0-prime condition (2.5 characters). The effect of semantic relatedness was only marginally significant ($F(1,17) = 3.47, \text{MSE} = 0.71, p < 0.10$).

As suggested in Experiment 1, the effects of semantic relatedness emerge more clearly in cases where the eyes are launched from close to the beginning of the target word. To test this hypothesis in Experiment 2, data were divided a posteriori in two classes depending on the launch site. Close launch sites corresponded to cases where the eyes were launched from a position 6 characters or less from the space preceding the beginning of the target word. Far launch sites corresponded to the remaining cases. The distinction between both launch site classes differed from that used in Experiment 1 (9 letters) and corresponded to the median of the launch sites found in the present experiment. Table 2 shows the mean initial landing sites in the target words as a function of both semantic relatedness and the launch site. It shows first that the eyes land further away from the beginning of the target word in close than far launch sites ($F(1,17) = 14.72, \text{MSE} = 13.30, p < 0.01$). This finding replicates the classical launch site effect (McConkie et al., 1988; Vitu et al., 1995). Second, in both 1- and 2-prime conditions, the eyes land further away from the beginning of the target word than in the 0-prime condition, but this holds only for close launch sites that favor a parafoveal processing of the target word. The effect of semantic relatedness was not significant, but there was a significant interaction between semantic relatedness and launch site ($F(1,17) = 7.66, \text{MSE} = 2.39, p < 0.01$).

There is no major difference between 1- vs. 2-prime conditions. Specific comparisons show that for close launch sites, both 2- and 1-prime conditions differ significantly from the 0-prime condition ($t(17) = 2.15, SD = 0.67, p < 0.05$ and $t(17) = 3.90, SD = 2.62, p < 0.01$, respectively) and that they differ also from each other ($t(17) = 2.67, SD = 3.46, p < 0.05$). In contrast, for far launch sites, both 2- and 1-prime conditions do not differ significantly from the 0-prime condition, or from each other.

In the present experiment, target words were again of variable lengths (between 5 and 7 letters), which (as was argued in Experiment 1) may account for the effect of semantic relatedness if the post-hoc partitioning of the data resulted in different word lengths in the different conditions. To ensure that this was not the case, initial landing sites in the target words were re-coded using the interval algorithm of
Experiment 1. The distributions of re-coded initial landing sites as a function of semantic relatedness for both launch sites are presented in Figs. 4(a) and (b). The figures show for close launch sites a clear shift of the initial landing sites towards the end of the words for 1- and 2-prime conditions in comparison with the 0-prime condition. In contrast, the three distributions overlap to a certain extent for far launch sites.

Analyses of variance on the re-coded data showed that the eyes land further away from the beginning of the target word for close than for far launch sites ($F(1, 17) = 23.20, \text{MSE} = 0.12, p < 0.01$). This finding again replicates the classical launch site effect (McConkie et al., 1988; Vitu et al., 1995). The effect of semantic relatedness was significant ($F(2, 34) = 3.29, \text{MSE} = 0.06, p < 0.05$), and the interaction between semantic relatedness and launch site was marginally significant ($F(1, 17) = 2.54, \text{MSE} = 0.07, p < 0.10$).

Specific comparisons on re-coded landing sites showed that for close launch sites, both 2- and 1-prime conditions differ from the 0-prime condition ($t(17) = 2.21, SD = 0.39, p < 0.05$ and $t(17) = 3.67, SD = 0.34, p < 0.01$) and that they did not differ from each other. For far launch sites, both 2- and 1-prime conditions did not differ significantly from the 0-prime condition, or from each other.

Fig. 4. Distributions of the initial eye landing positions in the target words (in word zones) as a function of the number of primes which are related to the target and a previously read prime (0,1,2), for close (a) and far launch sites (b).
Prior fixation durations. To ensure that the effects of semantic context in the present experiment did not result from differences in prior fixation durations, the mean duration of the fixation that preceded the initial fixation on the target word was calculated for 0-, 1-, and 2-prime conditions. The results are presented in Table 3. They indicate a tendency for prior fixation durations to be longer in the 0-prime condition than in both 1- and 2-prime conditions, and for far compared to close launch sites. However, an analysis of variance showed that none of the effects were significant.

Conclusion. Semantic context in the form of a related prime affects initial landing sites, provided that on the prior fixation, the eyes are close enough to the word for parafoveal preprocessing to occur. This finding replicates the findings of Experiment 1 with high-frequency words. However, adding a second prime did not clearly enhance this effect.

6. General discussion

To determine the role of ongoing perceptual and linguistic processes on inter-word saccades in reading, the present study examined the influence of semantic context on initial landing sites in target words. Both Experiments 1 and 2 showed a clear shift of the initial landing sites toward target words as a result of the presence of a preceding semantically related prime. The effect was found for high-frequency target words and when the preceding fixation was close to the beginning of the target word.

The present results contrast with previous findings showing that linguistic factors cannot affect, or only slightly affect, initial landing sites in words (Rayner & Morris, 1992). The findings are compatible, however, with several studies reporting an effect of semantic context on both initial landing sites (Dubois & Sprenger-Charolles, 1988) and word skipping rate (Ehrlich & Rayner, 1981; Balota et al., 1985; see also O’Regan, 1990; Brysbaert & Vitu, 1998). However, unlike some of these studies, the present effects cannot be accounted for by possible confounding variables such as the visual characteristics of the target word, the frequency of occurrence of the target word, the position where the eyes were launched from relative to the beginning of

<table>
<thead>
<tr>
<th>Number of related primes</th>
<th>Launch site</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Close</td>
</tr>
<tr>
<td>0</td>
<td>315</td>
</tr>
<tr>
<td>1</td>
<td>296</td>
</tr>
<tr>
<td>2</td>
<td>293</td>
</tr>
</tbody>
</table>

Table 3
Mean duration (ms) of the fixation that preceded the initial fixation on the target word as a function of the number of semantically related primes (0,1,2) and the launch site (close vs. far; Experiment 2)
the target word, or the prior fixation duration. Indeed the same high- and low-frequency target words were used in both related and unrelated prime conditions. In addition, the launch site was controlled a posteriori and it was shown that the duration of the fixation that preceded the target word did not vary with semantic relatedness.

The present study allows the identification of factors that are critical for finding semantic context effects on initial landing sites of eye fixations. An effect of semantic context can be observed only when the word is very frequent in the language and when the eyes are close enough to the word in order to be preprocessed in parafoveal vision. This finding argues in favor of the hypothesis that this is only when a number of perceptual and lexical constraints are combined and therefore when processing of the word in parafoveal vision is strongly facilitated that semantic context can affect where the eyes initially land in a word (Vitu, 1991a). Indeed, close launch sites amounted to about 6–8 letters from the space preceding the beginning of the target word. This is close to the limits of the perceptual span, which is a necessary condition for preprocessing of the word in parafoveal vision to occur (McConkie & Rayner, 1975; McClelland & O’Regan, 1981; Brysbaert et al., 1996). In addition, high-frequency words in either foveal or parafoveal vision are known to be more rapidly processed (McClelland & Rumelhart, 1981; Inhoff & Rayner, 1986; Vitu, 1991a).

At a theoretical level, the present results argue against the proposal made by the oculomotor-control view that ongoing perceptual and lexical processes associated with parafoveal word identification cannot intervene in the decision where to send the eyes next within an average fixation duration (O’Regan & Lévy-Schoen, 1987; O’Regan, 1990). Rather, the results confirm cognitive processing-type views of eye guidance in reading: the likelihood that initial landing sites in words change with linguistic factors is a function of the easiness of processing the word in parafoveal vision and is more generally dependent on the respective time course of saccade programming and parafoveal word identification processes (Morrison, 1984; Henderson & Ferreira, 1990; Reichle et al., 1998; Rayner, Reichle & Pollatsek, 1998; Rayner, Sereno & Raney, 1996). It must be noted, however, that this does not necessarily imply that ongoing linguistic processing is the main driving force in guiding eye movements. Low-level visuomotor constraints probably are the primary determinants of the next eye location, but in addition higher level processes are able to modify the aimed-for location. This is hard to show experimentally simply because several constraints must be combined for ongoing processing to be fast enough to affect the current saccade plan.

The following mechanisms underlying the observed effects of semantic context on saccade size may be conceived. After reading the prime, the system generates some predictions as to the upcoming target word which leads to a pre-activation of the target word. When the eyes get closer to the target word and some letter-information of the word can be extracted in parafoveal vision, then both types of information are combined and the target word will get even more activated if both types of information are compatible. This process will be enhanced if the word is very frequent, because high-frequency words are characterized with higher resting
levels of activation and/or higher activation rates (McClelland & Rumelhart, 1981),
and because the amount of parafoveal processing increases with the word frequency
(Inhoff & Rayner, 1986; Vitu, 1991b). However, if the last eye position before the
target word is not close enough for some letters of the word to be extracted in
parafoveal vision, and/or if the word is of a low frequency of occurrence, then the
activation level of the target word will be similar to that of other target words
generated by the prime.

When the system gets ready to move the eyes to the next word (which is assumed
by most authors to occur when the fixated word is identified), a saccade is prepared
which aims for the next word. The size of the saccade is calculated on the basis of a
low-level analysis of the visual configuration of the following words (Vitu, 1991a).
However, if the activation level of the target word reaches a given threshold before
the saccade is ready to be executed, then the saccade is lengthened either to skip the
target word if it is identified, or to land further into the target word. The lengthening
of the saccade brings the eyes at a position which favors the extraction of the target’s
letters that could not be extracted in parafoveal vision on the prior fixation and helps
speed up the processing of the target word. This also serves to bring the eyes closer to
the next word, and therefore helps start the preprocessing of this word while pro-
cessing of the target word is being achieved.

The present results suggest that context effects intervene very early in the time
course of word processing and before the word is lexically identified, within the 250
ms of parafoveal processing, and not only within the 500 ms of reaction time in
lexical decision or naming tasks. Prior studies have shown that the information
extracted from a word in parafoveal vision is incomplete (Lima & Inhoff, 1985). It
is therefore unlikely that the target words in the present experiments, which were
between 5 and 8 letters (6–8 in Experiment 1 and 5–7 in Experiment 2) could be
fully identified before the eyes land on them. Thus, it is unlikely that semantic
context effects on initial landing sites resulted from post-access strategies occurring
after the word was identified (VanVoorhis & Dark, 1995; Thompson-Schill et al.,
1998). Actually, if the target word had been fully identified in parafoveal vision,
then we might wonder why such predictable words were fixated at all and not
skipped. As was noted by Balota et al. (1985), words can be skipped even though
they have not been fully identified in parafoveal vision on a prior fixation, sug-
gestig that context effects on word skipping rate might also reflect an early locus
of context effects. In addition, the fact that semantic context interacts with both
word frequency (Experiment 1) and launch site (Experiments 1 and 2) provides
further evidence in favor of the hypothesis of early context effects, given that word
frequency and launch site affect the lexical access in parafoveal vision (Vitu,
1991a,b).

Results argue for the hypothesis that context effects lead to forward anticipations
by means of automatic spreading activation (Collins & Quillian, 1969; Collins &
Loftus, 1975; Anderson, 1983) or expectancy mechanisms (Posner & Snyder, 1975a,b;
Neely & Keefe, 1989) leading to an activation of the target word in memory before it
is actually perceived. The assumption that the contextual prime serves to make
predictions as to the upcoming target word is compatible with a context-dependent
view for word identification and with the hypothesis of an early locus of context effects (Tabossi, 1988a,b; Tabossi & Zardon, 1993).

Given the timing constraints associated with processing of the foveal word and the programming of a saccade, the range of time within which the predictability of the target word affects the size of the saccade is limited. The likelihood that such an effect occurs is thus higher when several types of information accumulate, such as when the eyes are close to the word for some information from the word to be extracted in parafovea (close launch site) and the word is of a high frequency of occurrence. In this case it is possible to observe anticipatory context effects occurring at early locus of lexical access.

Acknowledgements

This research was supported by a grant given to Frédéric Lavigne (from the French Ministry of Research) and by two grants given to Françoise Vitu for collaboration with Géry d’Ydewalle’s Laboratory of Experimental Psychology at the University of Leuven in Belgium (‘Human Capital and Mobility’ from the European Community, and ‘Tournesol’ from the French Ministry of Foreign Affairs).

We would like to express special thanks to members of the Laboratory of Experimental Psychology at the University of Leuven in Belgium, Andreas De Troy for setting up the experimental software and giving technical assistance in conducting this research, and Marc Brysbaert and Marlene Devijver for their help and encouragement. We also thank Keith Rayner, Gezinus Wolters, and an anonymous reviewer for their fruitful comments on a previous version of the article.

References


Trésor de la Langue Française (1971). Nancy: CNRS.


