

Eye movements as direct tests of the GO model for the missing-letter effect

ANNIE ROY-CHARLAND AND JEAN SAINT-AUBIN
Université de Moncton, Moncton, New Brunswick, Canada

AND

RAYMOND M. KLEIN AND MIKE LAWRENCE
Dalhousie University, Halifax, Nova Scotia, Canada

When asked to detect target letters while reading a text, participants miss more letters in frequently occurring function words than in less frequent content words. To account for this pattern of results, known as the missing-letter effect, Greenberg, Healy, Koriat, and Kreiner (2004) proposed the guidance–organization (GO) model, which integrates the two leading models of the missing-letter effect while incorporating innovative assumptions based on the literature on eye movements during reading. The GO model was evaluated by monitoring the eye movements of participants while they searched for a target letter in a continuous text display. Results revealed the usual missing-letter effect, and many empirical benchmark effects in eye movement literature were observed. However, contrary to the predictions of the GO model, response latencies were longer for function words than for content words. Alternative models are discussed that can accommodate both error and response latency data for the missing-letter effect.

In a letter detection task, in which participants are asked to read a continuous text and to circle all instances of a target letter, omissions are far more numerous when the target appears in frequent function words than when it appears in less frequent content words (Corcoran, 1966). This well-replicated phenomenon, known as the *missing-letter effect*, has benefited from intense empirical activity and a vigorous theoretical debate between the two leading proposals, the unitization account/processing time hypothesis and the structural account (Healy, 1994; Koriat & Greenberg, 1994). Although each of these models has received considerable empirical support, neither of them individually accounts for all results in the literature. As evidence accumulated, it became increasingly evident that an integrative model would provide a better explanation (Saint-Aubin & Klein, 2001; Saint-Aubin & Poirier, 1997). Recently, Greenberg, Healy, Koriat, and Kreiner (2004) put forward such an integrative proposal; called the *guidance–organization (GO) model*, it integrates the major assumptions of the two competing models, and also proposes innovative assumptions centered on eye movement patterns. The aim of the present study was to test the key assumptions of this promising new proposal by recording eye movements while participants engaged in a letter detection task.

The missing-letter effect is usually observed by comparing a very frequent function word with less frequent content words (Healy, 1994; Koriat & Greenberg, 1994).

However, as this description implies, two factors usually covary in this contrast: word function and word frequency (Koriat, Greenberg, & Goldshmid, 1991). Not surprisingly, many studies have been conducted to assess the unique contribution of each of these factors. For instance, the influence of word frequency was evidenced by a higher omission rate for the target letter embedded in more frequent content words than in less frequent content words (e.g., *time–tine*, *cost–cyst*, *thought–thicket*) (Greenberg, Koriat, & Vellutino, 1998; Healy, 1976; Minkoff & Raney, 2000; Roy-Charland & Saint-Aubin, 2006; Saint-Aubin & Klein, 2004; Saint-Aubin, Klein, & Landry, 2005), as well as in the frequent personal pronoun *they* compared with the rare and archaic personal pronoun *thou* (Moravcsik & Healy, 1995). The influence of word function has been shown, among other means, by a higher omission rate of the target letter *t* embedded in the function word *at* than in the content word *it*, despite the fact that both words occur with about equal frequency (Greenberg et al., 1998; Roy-Charland & Saint-Aubin, 2006; Saint-Aubin et al., 2005). Similarly, in French, Saint-Aubin and Poirier (1997) observed more omissions of the target letter *r* in *or* when used as a function word meaning *whereas* than as a content word meaning *gold*, even though both usages occur with about the same frequency.

Until very recently, the missing-letter effect was accounted for either by calling on processes mostly influenced by word frequency (unitization account/processing

time hypothesis) (Healy, 1994; Moravcsik & Healy, 1995) or on those mostly influenced by word function (structural account) (Koriat & Greenberg, 1994). As none of these models was able to account for the entire pattern of results, an integrative hypothesis, the GO model, was proposed by the leading proponents of the two competing models (Greenberg et al., 2004). This model is an important addition to the missing-letter effect literature, since it answers the major criticisms of previous models (Saint-Aubin & Klein, 2001). Within the GO model, the core components of the two previous models have been adopted as a set of five main assumptions. Unfortunately, as presented by Greenberg et al. (2004), it is unclear to what extent all the assumptions of the two root models are carried forward into the integrated framework. We will further develop this point below.

According to our understanding of the GO model, the proximal cause of the missing-letter effect is the time spent processing the target word. This assumption is taken directly from the unitization account/processing time hypothesis on which the GO model is based (Moravcsik & Healy, 1995). According to this view, readers process text in parallel at several levels of analysis: feature, letter, syllable, and word. It is assumed that the time required to process a given letter is constant, whatever the word is in which the letter is embedded. We further assume that a unit, like a word,

can be identified either on the basis of visual features or as the result of the identification of component lower-level units. . . . The important assumption of the unitization hypotheses, which differentiates them from the interactive-activation model, is that once a unit has been identified, the processing of its component lower-level units is terminated even if they have not yet reached the point of identification; instead, processing and attention are driven to the next location in the text. (Healy, 1994, p. 336)

From this viewpoint, word function and word frequency would be distally related to omission rate by influencing both word identification speed and, consequently, the time available for letter processing. High frequency words would be identified more rapidly as a whole at the word level, due to the higher familiarity of their visual configuration. It is further assumed that, even after allowing for the influence of word frequency, function words would be identified faster than content words because of the contextual constraints which help to identify sentential slots where structure-supporting function words may be found (Greenberg et al., 2004). Very early in sentence processing, to allow the integration of meaning-laden elements, readers strive to establish a tentative frame of the phrase or the entire sentence, so their attention would rapidly shift from the structural units; consequently, there would be less time available for letter processing of function words. In addition, because they are monitored on the basis of a shallow and rapid analysis guided by contextual constraints, function units are frequently identified in the parafovea; and it is further assumed within the GO model

that letters could not be identified in the parafovea, because the visual acuity there is too low. Consequently, it would be impossible to detect a target letter if the word were skipped. Despite the clarity of processing time assumption, it is worth mentioning that the GO model does not provide an explicit mechanism for the detection of target letters.

Although the GO model has not yet been systematically tested, certain findings in the literature are pertinent to some of its assumptions. As mentioned above, omissions within the GO model are due to faster processing at the word level, which would often terminate letter processing before the target letter had been identified. This translates into higher omission rates for targets embedded in frequent words than for those embedded in rare words, and in function words than in content words. Importantly, as explained by Saint-Aubin, Klein, and Roy-Charland (2003), and as can be seen in the upper left panel of Figure 1, this unitization assumption predicts that higher omission rates will be associated with more rapid responses to the letter targets. The reason is that target response latencies come from a common distribution of letter processing finishing times that is truncated due to interruption of the search for the target letter (in the target-containing word) when word processing is completed before the target is identified. More precisely, it is the interruption of the letter processing when a word is identified that would cause the response latency distribution to be truncated. Higher rates of omission, then, would be associated with the slower letter finishing times being more greatly truncated; therefore, faster letter-target response latencies would be observed. In order to make a direct test of this assumption, response latencies had to be measured. Although it is not possible to do so with the usual paper-and-pencil procedure, Saint-Aubin et al. (2003) accomplished it using a rapid serial visual presentation (RSVP) procedure. In this procedure, words are presented one by one on a computer screen; while participants are studying the passage, they are asked to press a response key as soon as they detect a target letter. In a series of five experiments, Saint-Aubin et al. (2003) observed more omissions for function words than for content words and, among content words, more omissions for more frequent words. Most importantly, and contrary to the prediction of the GO model (see the upper left panel of Figure 1), response latencies were longer for targets embedded in function words than in content words and in frequent words than in rare words.

Although unequivocally contrary to the predictions of the GO model, it has been argued that response latency results obtained with an RSVP procedure are not the most critical tests of the unitization assumption, because word presentation duration is fixed, not under the reader's control.¹ Saint-Aubin and Klein (2004) addressed this issue when they compared omission rates for the same text using two procedures: a paper-and-pencil procedure and an RSVP procedure. They found that the correlation between both procedures was as high as the reliability estimate of each measure. Saint-Aubin and Klein (2004) clearly established the reliability of the RSVP procedure for investigating the

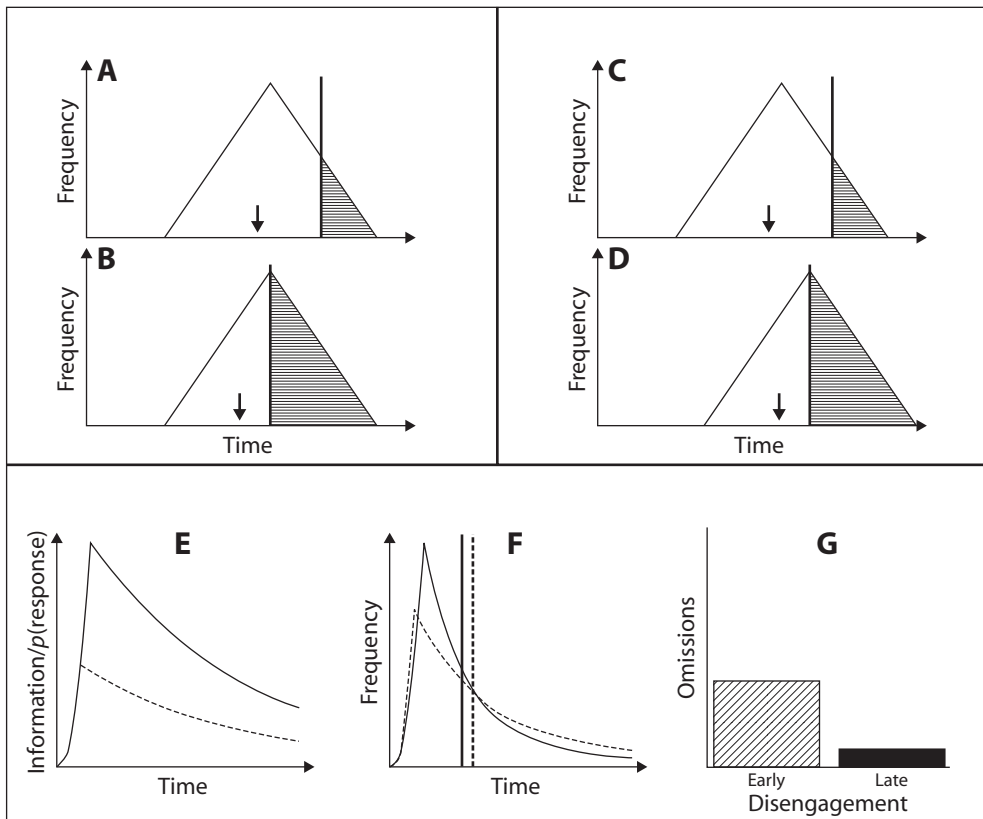


Figure 1. Three models of the missing-letter effect: the guidance–organization (GO) model (A and B), an alternative model proposed by Saint-Aubin et al. (2003) (C and D), and the attentional-disengagement (AD) model (E, F, and G). In A and B, the finishing times of target detection are shown as a triangular distribution in A or a high frequency function word in B. The horizontal position of the gray vertical line represents the average time to fully process the critical word in which the target letter is embedded. The vertical line is more toward the left on the x-axis in B, representing high frequency function words, than in A, representing lower frequency content words. The downward pointing arrow represents the mean target detection time on trials when the target was detected. The shaded portion of the triangular distribution represents omissions which occurred because the reader would have stopped processing the target letter before it was fully processed. In C and D, illustrating the alternative model proposed by Saint-Aubin et al. (2003), mean response latencies are indicated by downward pointing arrows and omission of targets by the shaded regions. At the bottom, E, F, and G illustrate our AD model, which emphasizes the role of attention during reading. The solid curve in E represents a content word and the dotted curve represents a high-frequency function word; the peak of both curves represents the moment at which the reader disengages attention from the critical word in which the target letter is embedded. The shape of the curve before the peak is exactly the same for content and high-frequency function words. The peak is higher for content words because attention remains on the critical word for a longer amount of time. The curve after the peak represents decay. The probability of response at a given point in time is directly proportional to the amount of information accumulated at that time. Monte Carlo simulations with 1,000,000 iterations, a resolution of 1 msec, and a response timeout of 2,200 msec produced the data shown in F and G; F shows the means (vertical lines) and kernel smoothed density distributions of response latencies for the late (content words; solid lines) and early (high-frequency function words; dotted lines) disengagement times; G shows omission rates from the same simulation. Code for this simulation in R is available by request.

missing-letter effect when omission rate is the measure of primary interest. However, it remains to be seen if the same similarity among results can be observed when response latencies are the measures of primary interest; in other words, would the pattern of response latency be the same when the presentation rate is constrained, as in the RSVP procedure, as when the presentation rate is under the reader's control, as in normal reading? This could be determined by using a

letter detection task while monitoring eye movements, as in the work of Saint-Aubin and Klein (2001)—who, unfortunately, did not compute response latencies.

Eye movement monitoring will also allow the implementation of a test of an implicit assumption of the GO model. More precisely, by extensively calling upon results of previous studies monitoring eye movements to explain the missing-letter effect, the authors argue that the typical

patterns of eye movements reported in the literature are unaffected by the search task that must be performed in parallel to the reading task in order to study the missing-letter effect. On this matter, an extensive discussion, not based on direct evidence from eye movements, can be found in Healy's (1994) presentation of the unitization account/processing time hypothesis on which the GO model is partly based. Similarly, Greenberg and Tai (2001) suggested—without testing their hypothesis—that the disappearance of the missing-letter effect in the critical conditions of their study was due to a distorted pattern of eye movements. More recently and more precisely, Greenberg et al. (2004) called upon eye movement data to support their main assumptions, and in many cases they suggested that the cognitive processes they proposed are mediated by eye movements. For instance, word frequency and item predictability are assumed to influence viewing time as measured by gaze durations (which would influence the time available for letter processing), because it is assumed that letter processing is most likely to occur when the word is fixated. It is further assumed that word frequency and word function influence the skipping rate, with high-frequency and function words being more likely than low-frequency and content words to be skipped in reading. Because skipped words during reading are assumed to be identified in the parafovea when the previous word is fixated, and because it is well known that the quality of the visual input decreases as a function of the distance between the target and the fixation location, it is assumed that word function and word frequency would influence letter detection accuracy by influencing the quality of the visual input.

Although Greenberg et al. (2004) quite accurately reviewed previous studies in the eye movement domain, it remains to be determined whether or not the critical findings actually generalize to the letter search task paradigm. In effect, although interesting, there are doubts about the possibility of extending the eye movement database to a situation in which a letter search task is performed simultaneously with reading. More precisely, after reviewing the literature on eye movements when participants are searching for a target in a text, Rayner and Pollatsek (1989) concluded that the search task requirement may destroy the usual pattern of eye movements. For instance, Rayner and Raney (1996) conducted a study in which participants were asked to either read normally or search for a target word (see Spragins, Lefton, & Fisher, 1976). Critical word pairs were matched for meaning while varying in frequency as a frequent content word was paired with a less frequent synonym (e.g., *flowers–blossoms*, *ancient–archaic*, *sharp–stark*). Results showed a higher skip rate for more frequent words than for less frequent words in the reading task, but no significant difference was found in the word search task. It must be noted, however, that Rayner and Raney's methods depart in several respects from the search task used in the missing-letter effect paradigm. In Rayner and Raney's study, participants in the search task condition were scanning the text for a target word, not a target letter; and, although in the

missing-letter effect literature the typical task is to read for comprehension and search for a target letter, the effect has been observed when participants were searching for target words as well (Saint-Aubin et al., 2003). However, in studies of the missing-letter effect, the words in which the target letter is embedded, or those for which participants are searching, are usually the focus of observation. For example, Saint-Aubin et al. (2003) had participants searching for either the word *pour* (for) or the word *cour* (yard). They observed response latencies and omission rate for these target words. In Rayner and Raney's study, participants were not searching for the words from the matched pairs for which eye movement measures were observed; in fact, they were searching for another word. Moreover, the omission rate on the target word in Rayner and Raney's study was much lower than in the results observed by Saint-Aubin et al. (2003), when participants were reading for comprehension while searching for a target word. This highlights another difference in the tasks: In Rayner and Raney's study, participants were scanning the text in a left to right motion, to mimic reading without processing the meaning. In a letter search task typically used to study the missing-letter effect, participants read for comprehension while performing a search task; text comprehension is then assessed. This difference is critical, because it is the intention of most researchers of the missing-letter effect to generate potentially useful data on the nature of normal reading. Once the emphasis on comprehension is dropped, there is no guarantee that any such data will be generated.

The main purpose of the experiment reported here was to implement the first direct tests of the central assumptions of the GO model. The GO model was evaluated by contrasting a highly frequent French function word (*des*, a plural indefinite article) with three-letter control content words. The text used here was selected because it allows (1) the establishment of the baseline effect; (2) a test of the specific hypotheses regarding word frequency and word function effects; and (3) the establishment of the pattern, with the RSVP procedure, of response latencies, a task for which it has previously been used (Saint-Aubin & Klein, 2004; Saint-Aubin et al., 2003). The GO model predicts that, overall, there should be more omissions for the frequent function word *des* than for the less frequent control content words, and that there should be more omissions when the critical words are skipped than when they are fixated (as previously observed by Saint-Aubin & Klein, 2001). The higher omission rate for skipped words could be due to the lower visual acuity in the parafovea. It is worth noting that the GO model does not take into account the launch site, which is the position of the last fixation before the eyes cross the critical word. Within the GO model, if a word is skipped, target letter detection would almost be impossible, despite the well-known fact that visual acuity could be very good for the target letter—especially here, with the target letter as the first letter of the word—when the launch site is close to the critical word (Rayner, 1998). In addition, according to the GO model when the critical words are fixated, the gaze durations should be shorter for omissions than for

detections. This follows from the direct link between lexical processing time and fixation duration (Rayner, Ashby, Pollatsek, & Reichle, 2004; Reichle, Pollatsek, Fisher, & Rayner, 1998) and the GO model's assumption that when the word is processed faster there is less time available for letter processing, and that the probability of an omission is, therefore, higher. There should also be a positive correlation between the log frequency of the content words and their omission rate.

Taking into account the architecture of the GO model, and on the basis of our knowledge of eye movements, we derived a series of predictions about response latencies—defined as the time elapsed between the moment the eyes cross an invisible boundary, located on the first letter of the target word, and the moment the response button is pressed. First, response latencies should be longer when the critical words are skipped than when they are fixated; this would reflect the fact that it is harder to detect a target letter in the parafovea, where visual acuity is lower. Second, excluding the possibility of a target being detected after a regressive saccade, response latencies should be shorter for frequent words than for rare ones, and for function words than for content words, for the reasons given above (see also the upper left panel of Figure 1 and Saint-Aubin et al., 2003). For targets detected after a regressive saccade, the same prediction would be made, for much the same reasons. It is important to consider, however, that target detection response latencies would necessarily be slower when detection is enabled by a regressive saccade, and that, therefore, overall target response latencies may depend on the relative frequency with which such regressions occur for the different target conditions.

Finally, the GO model also acknowledges variations in the processing of different occurrences of the same function word. It is assumed that “contextual constraints are critical in helping to identify sentential slots where structure-supporting units are likely to reside. Such constraints operate in a top-down manner to assess whether a parafoveal pattern is signalling the start of a new unit” (Greenberg et al., 2004, pp. 430–431). In support of this hypothesis, Koriat and Greenberg (1991) observed more omissions when a nonword was inserted in a function slot than in a content slot in a sentence (see Healy, 1994, for an alternative interpretation). In the present study, a systematic examination of this hypothesis was implemented by collecting probability norms for each occurrence of the critical words. First, norms were collected in regards to the proportion of participants expecting a function word, be it the target word or not. According to the GO model, if contextual constraints signal a sentential slot where a function word would reside, it is more likely that the word would be skipped: The more participants are expecting a function word, therefore, the higher the skip rate should be. Second, norms for proportion of accurate predictions—that is, the proportion of participants correctly predicting the target word—were also gathered in order to provide a direct comparison with the previous studies in the eye movement database in which a higher skip rate was observed for more predictable words (see Balota, Pollatsek,

& Rayner, 1985). These previous results are also in line with the predictions of the GO model. In fact, the higher the predictability of a word, the faster its identification in the parafovea, or when fixated, which would translate into higher skip and omission rates.

METHOD

Participants

Forty-two unpaid undergraduate students (33 women, 9 men) at the Université de Moncton participated in the experiment. All participants had normal or corrected-to-normal vision, and French was their first language.

Materials

The *des* text used by Saint-Aubin and Klein (2004) and by Saint-Aubin et al. (2003) was used here. It contained 593 words and 24 instances of the French plural indefinite article *des*, with a frequency count of 15,877 occurrences per million; 24 instances of three-letter control content words beginning with the target letter *d*, with an average frequency count of 248 occurrences per million—for example, *don* (gift or donation), *dit* (says), *dix* (ten), *dos* (back), and *duo* (duet) (New, Pallier, Ferrand, & Matos, 2001). In addition, the target letter *d* was also embedded in 14 noncritical words varying in length from 2 to 10 letters. Each word containing the target letter, critical or not, was separated from the previous and subsequent ones by at least four filler words without target letters. The test words were not included in the first and the last sentence of the text; they were never at the beginning or the end of a sentence or a line; and they were never adjacent to a punctuation mark.

Apparatus for Eye Movement Monitoring

Eye movements were measured with an SR Research, Ltd., EyeLink II system. This system has high accuracy ($<0.5^\circ$) and a high sampling rate (500 Hz). The EyeLink headband has three cameras, allowing simultaneous tracking of both eyes and head position to compensate for head motion. By default, only the pupil (without corneal reflection) of the participant's dominant eye was tracked in our study. The EyeLink II system uses an Ethernet link between the eyetracker and the display computer for real-time saccade and gaze-position data transfer. In the present investigation, the configurable acceleration and velocity thresholds were set to detect saccades of 0.5° or greater.

Stimulus displays were presented on two monitors, one for the participant (19-in. ViewSonic monitor) and the other for the experimenter. The experimenter's monitor was used to give feedback in real time about the participant's computed gaze position. This feedback was given in the form of a gaze cursor measuring 1° in diameter, allowing the experimenter to evaluate system accuracy and to initiate a recalibration if necessary.

Procedure

Participants took part individually in a private room, in a single session lasting approximately 10 min, of which the first 5 min were used to calibrate the eyetracking system and the last 5 min for the reading task. Participants were seated about 60 cm from the screen. They were told that they would see a four-page prose passage on the computer screen. The text was presented double spaced, in the Times New Roman font in black on a white screen. Between 11 and 13 lines were presented per page. At 60 cm from the screen, a horizontal character averaged 0.44° of visual angle, and the vertical spacing between lines equaled 1° of visual angle. Each page presentation was separated by a short pause that was used to verify the calibration of the system. The participants were instructed to read the text for comprehension at their normal reading rate and to search for the target letter *d*. Participants signaled target letter detection by clicking the mouse as soon as they saw a target letter. To promote

reading for comprehension, participants were also informed that they would have to answer three multiple-choice comprehension questions on a sheet of paper after reading the entire text displayed on the four screens.

RESULTS

The data for the reading and searching tasks were scored with the EyeLink Data Viewer program, which presented the text with participants' successive landing positions superimposed on it. First, for each of the 48 critical words (24 *des* and 24 control content words), omissions were computed by means of the Data Viewer, which displayed as an event each participant's buttonpress response. For each event, both the position of the eye when the response occurred and the time elapsed since the beginning of the screen presentation were displayed. To avoid any ambiguous attributions, a buttonpress had to occur before the participant's eye reached another word containing the target letter. Second, for the first set of analyses, a word was considered fixated, whether through the first pass or through a regression, when at least one landing position was on the word itself. These first two criteria are the same as those of the only study to date (Saint-Aubin & Klein, 2001) in which eye movements were monitored when participants performed a letter detection task while reading. The criteria were applied unaltered for the sake of comparability. In addition, for detected targets, response latencies were computed. (Response latencies represent the time elapsed from the moment the participant's eye crosses an invisible boundary at the first letter of the critical words to the moment the participant presses the response button.) More precisely, during the saccade, the timer was started as the eye crossed the letter *d* of the target words and was stopped when the button was pressed. With this criterion, whether the critical word was fixated or not, it was possible to compute response latencies. The probabilities of fixation results are presented first, followed by omissions, response latencies, and fixation durations.

Probability of Fixation

Proportions of fixated target words were computed for each participant as a function of test word condition (*des* vs. control content words) by dividing the number of fix-

ated target words by the number of occurrences. For the first analysis, as done by Saint-Aubin and Klein (2001), a word was considered fixated, whether in the first pass or during a regression. In all analyses, the .05 level of significance was adopted. Results of the first analysis revealed that the probability of fixating the function word *des* (.75) was significantly smaller than the probability of fixating the control content words (.82) [$F(1,41) = 6.37$, $MS_e = 0.018$]. A further analysis was computed for trials in which the word was fixated without a regression. Results revealed that the probability of fixating the function word *des* (.67) was again significantly smaller than the probability of fixating the control content words (.78) [$F(1,41) = 9.19$, $MS_e = 0.027$].

Omissions

Proportions of omissions for the target words are presented in Table 1 as a function of fixation status. These proportions were computed for each participant and each condition (2 test word \times 3 fixation status). The number of occurrences of target words in each condition varied from participant to participant because of variations in the number of fixated words. Furthermore, the analyses were based on 36 of the 42 participants, because 3 participants made no regression on any content words, 2 made no regression on any occurrences of the function word, and 1 did not skip any occurrences of the function word. Inspection of Table 1 reveals that the proportion of omissions was higher when the target letter was embedded in the function word *des* than when it was in the control content words, and higher when the critical words were skipped rather than fixated; also, no obvious difference in omission rate was observed between words fixated with or without a regression. These trends were confirmed by a 2×3 repeated measures ANOVA with test word (*des* and control content words) and fixation status (fixated with a regression, fixated without a regression, and skipped) as factors. Results revealed a significant main effect of test words [$F(1,35) = 18.39$, $MS_e = 0.069$] and of fixation status [$F(2,70) = 19.56$, $MS_e = 0.057$]; the interaction was not significant ($F < 1$). Post hoc comparisons (Tukey, HSD) revealed that there were more omissions when the critical word was skipped than when it was fixated, and that there was no difference between a fixation with a regression and one without.

Table 1
Means, Standard Deviations, and the Number of Observations per Cell for Probability of Omissions of the Target Letter and Response Latencies (RTs) As a Function of Fixation Status and Test Words

Target Type	Fixation Status											
	Without Regression				With Regression				Skipped			
	RT (<i>N</i> = 25)		Omissions (<i>N</i> = 36)		RT (<i>N</i> = 25)		Omissions (<i>N</i> = 36)		RT (<i>N</i> = 25)		Omissions (<i>N</i> = 36)	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Function	803	176	.32	.26	948	161	.27	.28	847	219	.49	.40
<i>n</i>	217		424		123		222		97		210	
Content	675	123	.14	.11	875	168	.12	.20	666	224	.36	.31
<i>n</i>	329		534		100		168		74		151	

Response Latencies

Letter detection response latencies were computed by averaging response latencies of detected target letters for each condition and participant. The analysis was based on 25 of the 42 participants, because when the words were skipped, 2 participants made no detection of any of the target words, 7 made no detection of the target letter embedded in the function word *des*, and 1 made no detection of the target letter embedded in the content words. When a regression occurred, 1 participant made no detection of any occurrences of the function word. Also omitted were the 6 participants who made no regression on any content words, or on any occurrences or skips on occurrences of the function word. Note that the longest response latency confidently attributed to a test word was 2,200 msec. As shown in Table 1, response latencies were longer when the target letter was embedded in *des* than when it was in the control content words, and when the word was fixated through a regression, than when it was skipped or fixated without a regression. No obvious difference was observed between the latter two. The 2×3 repeated measures ANOVA with test word and fixation status confirmed a main effect of test word [$F(1,24) = 36.20$, $MS_e = 16,786$] and of fixation status [$F(2,48) = 21.47$, $MS_e = 20,999$]; the interaction was not significant ($F = 1.67$, $p = .20$). Post hoc comparisons (Tukey, HSD) revealed that response latencies were longer when there was a regression than when the critical word was either skipped or fixated without a regression. There was no difference between the latter two conditions. In addition, to further test the reliability of response latency results, means were computed on all participants, instead of only those included in the ANOVAs. As mentioned previously, some participants were excluded from the analyses because of missing observations in one of the conditions (e.g., no omission for content or function words; no detection for content or function words). Even after increasing the number of observations per cell by using all available observations, the pattern of results remained the same.

Fixation Duration Measures

The most commonly used eye movement measures of fixation duration were applied to the data set (Rayner, 1998; Rayner et al., 2004). Specifically, the following

measures were used for each participant and each target word: (1) first fixation duration (the duration of the first forward fixation on a target word independent of the number of fixations made on the word); (2) single fixation duration (the duration of the fixation when only a single forward fixation was made on the target word); (3) gaze duration (the sum of all forward fixations on a target word prior to a saccade to another word); and (4) total fixation time (the sum of all fixations, including regressions and forward fixations if there was a regression on the target word). Measures 1–3 are conditional on the target word being fixated in the initial pass through the text, whereas measure 4 included all fixations made on a word, whether it was made in the initial pass or not.

Inspection of Table 2 reveals that all fixation duration measures for both function and content words were shorter when an omission occurred than when a detection occurred. Furthermore, forward fixation durations were longer for content than for function words, but only when an omission occurred for total fixation time. These trends were analyzed with 2×2 repeated measures ANOVAs with test word (*des* vs. control content words) and detection status (detected vs. omitted) as factors for each of the fixation duration measures. For the first fixation duration and the gaze duration, 33 of the 42 participants were used, because 6 participants made no omission on content words, 1 participant made no omission on the function word, 1 participant made no forward fixation on omitted content words, and 1 participant made no forward fixation on omitted function words. For the single fixation duration, 30 of the 42 participants were used; the same 9 participants that had been omitted from the analysis for the first fixation duration and gaze duration were omitted again. Furthermore, 2 participants always made multiple fixations on omitted content words and 1 participant always made multiple fixations on detected function words. For all of the forward fixation durations—that is, first fixation duration, gaze duration, and single fixation duration, respectively—the ANOVAs revealed a main effect of detection status [$F(1,32) = 5.78$, $MS_e = 3,598$; $F(1,32) = 10.39$, $MS_e = 4,599$; $F(1,29) = 10.44$, $MS_e = 4,259$] and a main effect of role [$F(1,32) = 6.41$, $MS_e = 3,138$; $F(1,32) = 6.57$, $MS_e = 3,832$; $F(1,29) = 5.59$, $MS_e = 3,129$], but the interaction did not reach significance ($F = 1.59$, $p = .22$; $F = 3.67$, $p = .07$; $F = 1.85$, $p = .18$). Planned contrasts revealed that the forward fixation durations were shorter on *des* than on the control content words when an omission occurred with, respectively, first, gaze, and single fixation durations [$F(1,64) = 7.32$, $MS_e = 2,969$; $F(1,64) = 10.19$, $MS_e = 3,347$; $F(1,58) = 7.19$, $MS_e = 2,711$]. However, there was no difference between *des* and the control words when a detection occurred (all F s < 1).

Finally, for the total fixation time, 35 of the 42 participants were used, because 6 participants made no omission on content words, and 1 participant made no omission on the function word. The ANOVA revealed a main effect of detection status [$F(1,34) = 23.32$, $MS_e = 5,593$] and an interaction [$F(1,34) = 12.10$, $MS_e = 2,190$]; the main effect of role did not reach significance [$F = 2.69$, $p =$

Table 2
Means and Standard Deviations for
Fixation Duration Measures As a Function
of Target Detection Status and Test Word

	First Fixation Duration ($N = 33$)		Single Fixation Duration ($N = 30$)		Gaze Duration ($N = 33$)		Total Fixation Time ($N = 35$)	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Omission								
Function	217	56	218	53	223	61	249	70
Content	253	69	254	75	269	71	294	65
Detection								
Function	254	49	268	52	279	66	337	107
Content	267	47	280	43	289	54	327	63

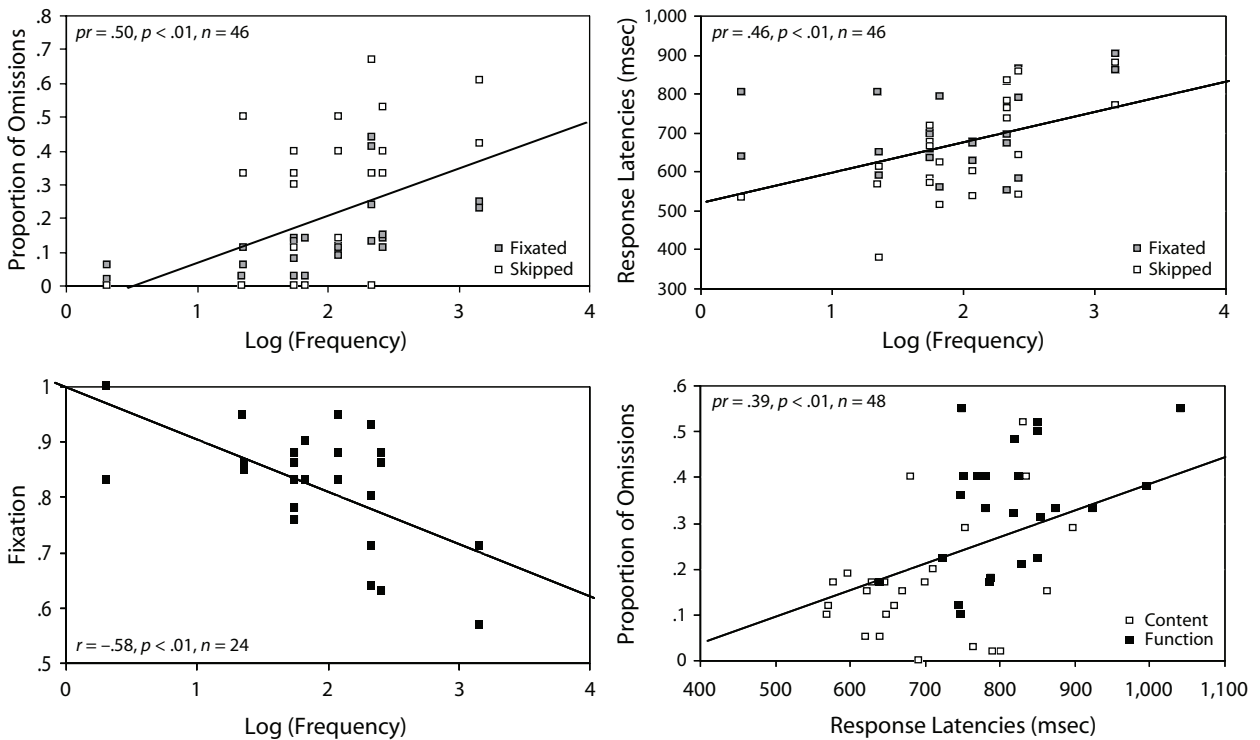


Figure 2. Log frequency value of content words as a function of omission rate (upper left panel) and response latencies (upper right panel) while controlling for fixation status, log frequency value of content words as a function of fixation status (lower left panel) as well as proportion of omissions as a function of response latencies while controlling for word class (lower right panel). The regression lines (upper panels) have been derived from the partial correlations of log frequency and the proportion of omissions and response latencies, while controlling for fixation status. The regression line (lower right panel) has been derived from the partial correlation between proportion of omissions and response latencies while controlling for word class.

.11]. Simple main effects revealed that the total fixation time was shorter for *des* than for the control content words when an omission occurred [$F(1,68) = 11.44, MS_e = 3,116$], however, not when a detection occurred ($F < 1$).

Word Frequency

Item-based correlations were computed in order to test the influence of word frequency on the probability of fixating a word for the 24 occurrences of the content words. Those 24 occurrences were composed of 9 different words, of which there were between 1 and 5 occurrences in the text. The analyses used the normal log of frequencies, because the frequency values varied between 2 and 1,462 occurrences per million (New et al., 2001). While controlling for fixation status, partial correlations were also computed between log frequency and omission rate as well as between log frequency and response latencies. Results for the correlation and partial correlations are presented in Figure 2. As can be seen in the upper panels of Figure 2, there were significant positive partial correlations between log frequency and omission rate as well as between log frequency and response latencies when controlling for fixation status. Furthermore, as can be seen in the lower left panel of Figure 2, a significant negative correlation was observed between log frequency and the probability of fixating a word. Finally, for the 48 target word tokens, partial item-based correlations were com-

puted between the proportion of omissions and response latencies while controlling for test words (*des* vs. control content words). As can be seen in the lower right panel of Figure 2, there was a significant positive partial correlation between the proportion of omissions and response latencies when a test word was controlled for.

Predictability

For the 48 target word tokens, item-based correlations were computed between normative data for the proportion of accurate prediction and the proportion of participants expecting a function word on the one hand, and skip rate, omission rate, and response latencies on the other. Partial correlations were also computed between the same measures while controlling for the test word role (function vs. content). The norms were established in a distinct experiment.² Results for the correlations and partial correlations are presented in Figure 3, which shows significant positive correlations between both predictability measures (item and function word) and skip rate, omission rate, and response latencies. The pattern of results is different for partial correlations. As can be seen in the upper left panel of Figure 3, there was a significant positive partial correlation between the proportion of accurate prediction and the skip rate. However, the middle and lower left panels show that partial correlations between proportions of accurate prediction and omission rate and response latencies were

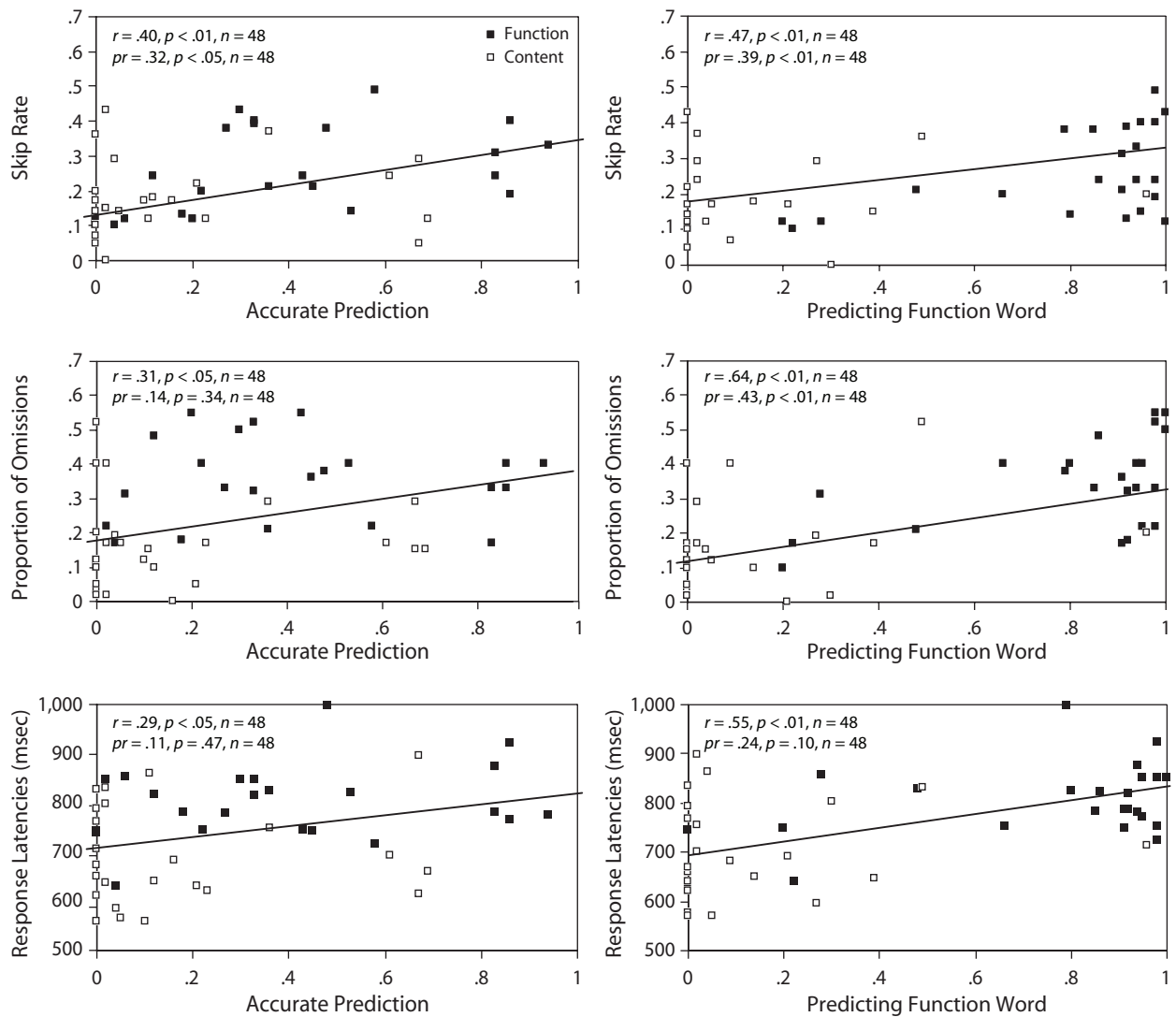


Figure 3. Proportion of accurate prediction of test words (left panels) and proportion of participants expecting a function word (right panels) as a function of skip rate (upper panels), omission rate (middle panels), and response latencies (lower panels), as well as partial correlations for these same measures while controlling for test word role. The regression lines have been derived from the total sample analyses without controlling for test word role.

not significant. Finally, the right panels of Figure 3 show significant partial correlations between the proportion of participants expecting a function word and both skip and omission rates, but not with response latencies.

DISCUSSION

The present results revealed more omissions and slower response latencies for the frequent function word *des* than for less frequent control content words, as well as a significant positive correlation for content words between their log frequency on the one hand and their omission rate and response latency on the other. As can be seen in Table 3, the magnitude of the missing-letter effect observed here is remarkably similar to that found with the same text, using a paper-and-pencil procedure (Saint-Aubin & Klein, 2004) and using an RSVP procedure (Saint-Aubin et al., 2003).

Not only does this suggest that our procedure for recording and coding detections and omissions was sound, it also points to the robustness of the missing-letter effect when confronted with very different methods of being studied. Second, as is also shown in Table 3, results for response latencies are also very similar to those found with the same text and an RSVP procedure (Saint-Aubin et al., 2003). Response latencies were longer for the detection of the target letter embedded in the function word *des* than for the detection of same in the control content words, and longer for frequent than for rare content words.

The results also reveal that there were more omissions when the target word was skipped than when it was fixated, but that the omission rate was unaffected by the presence or absence of a regression. First, our results perfectly mirrored, in French, those observed in English in the previous published study, in which eye movements were monitored

Table 3
Means and Standard Deviations for Probabilities of Omissions of the
Target Letter and Response Latencies From the Present Study,
and From Saint-Aubin and Klein (2004) and Saint-Aubin et al. (2003)

Target Type	Paper and Pencil (Saint-Aubin & Klein, 2004)		Eye Movement Monitoring (Present Study) (Overall, Fixated With and Without Regression, and Skipped)				RSVP (Saint-Aubin et al., 2003, Experiment 1)				Self-Paced RSVP (Saint-Aubin et al., 2003, Experiment 4)			
	Omissions (<i>N</i> = 68)		Omissions (<i>N</i> = 42)		RT (<i>N</i> = 42)		Omissions (<i>N</i> = 60)		RT (<i>N</i> = 60)		Omissions (<i>N</i> = 36)		RT (<i>N</i> = 35)	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Function	.34	.25	.33	.24	830	166	.41	.16	568	54	.54	.20	684	82
Content	.17	.14	.17	.11	699	129	.24	.14	546	44	.41	.19	650	80

while participants were reading for comprehension and searching for a target letter (Saint-Aubin & Klein, 2001). In addition, our detailed analysis of performance as a function of a regression extends the previous findings by revealing that regression patterns are inconsequential with regards to omission rate.

Eye movement results revealed that the skipping rate was higher for the frequent function word *des* than for the less frequent three-letter control content words. This result reproduced both the previous finding of Saint-Aubin and Klein (2001), with the letter search task performed parallel to the reading task, and the previous findings in the literature (Carpenter & Just, 1983; O'Regan, 1979; Rayner & Pollatsek, 1987). In addition, more predictable words were also more likely to be skipped than less predictable words. This finding also replicates previous results in the literature (Balota et al., 1985; Drieghe, Rayner, & Pollatsek, 2005). Not only do our results replicate previous findings, they also extend them. In effect, highly predictable words are usually compared with highly unpredictable words (e.g., "since the wedding was today, the baker rushed the wedding *cake/pies* to the reception"—Balota et al., 1985). Here, by using tokens of various predictabilities, we have shown that there is a monotonic relationship between predictability and skip rate. Finally, we also observed that as word frequency increases, so does skip rate. This finding also replicates previous results observed when participants are reading only, not also searching for a target letter (Raney & Rayner, 1995; Rayner & Raney, 1996).

With regard to fixation durations, results revealed that fixations were shorter with an omission than with a detection. Also, content words were fixated for a longer amount of time than was the function word *des*. More precisely, fixation durations were shorter for *des* than they were for content words for the subset of occurrences on which there was an omission. Results with omissions replicate previous findings for skipping rates obtained when participants were reading only (Carpenter & Just, 1983).

The GO Model

Most importantly, the present study provides the first systematic test of the major assumptions of the GO model for the missing-letter effect (Greenberg et al., 2004). More specifically, as described above, the GO model constitutes

an integrative hypothesis of the previous models of the missing-letter effect: the unitization account/processing time hypothesis (Healy, 1994; Moravcsik & Healy, 1995) and the structural account (Koriat & Greenberg, 1994). Furthermore, it also proposes innovative assumptions based on eye movement patterns.

The GO model accounts in a straightforward manner for the omission results. The omission rate was higher for *des* than for the three-letter control content words, and higher for the most frequent content words than for the less frequent ones. Processing of the function word *des* is faster both because it is more frequent than the content words and because the contextual constraints guide its identification. A strong correlation between the probability of expecting a function word and both skip and omission rates supported this view. The present results provide the first direct test of this assumption of the GO model, and they confirm the hypothesis that the higher omission of the target letter embedded in a nonword in a function slot instead of a content slot was due to the contextual constraints (Koriat & Greenberg, 1991). Finally, the higher omission rate for skipped than for fixated words would simply reflect the fact that visual acuity in the parafovea is lower, thus impeding letter identification (Hadley & Healy, 1991).

The pattern of fixations also fits nicely with the GO model. The shorter first, gaze, single, and total fixation durations for *des* than for its control content words that were observed when there was an omission would be due to the faster processing of *des*, because of both its higher frequency and its contribution to the structure of the sentence. Similarly, the higher skip rate for *des* than for its control words and for the most frequent content words than for the less frequent ones would also reflect the faster processing of the former. These results are interesting because they support the predictions of the GO model and replicate previous findings in eye movement monitoring studies when no search task was performed in parallel (Rayner, 1998).

In addition to the above results, we derived from the GO model a series of predictions about eye movements for which there were no previous data, predictions involving interactions between the pattern of detection at the search task and the pattern of fixations. More precisely,

the GO model makes more differential predictions about gaze durations when the target letter is detected than when it is omitted. Because there was only one previous study that combined the letter search task and eye movement monitoring, and because Saint-Aubin and Klein (2001) did not measure fixation durations or response latencies, it is easy to understand the lack of previous data. The predictions derived from the GO model were remarkably well supported. According to the GO model, an omission is produced when the word is processed faster than the target letter embedded in it. Because within the GO model gaze durations can be considered an index of word processing duration, shorter gaze durations were anticipated when the target letter was omitted than when it was detected. The results were exactly as predicted, and the same pattern was found with first, single, gaze, and total fixation duration measures. This pattern of results is even more striking in light of the fact that fixation durations on the word itself are not a pure index of word processing duration; for example, processing of word n can start when readers are fixating word $n-1$, which is known as the parafoveal preview benefit, and this can continue after the eyes have landed on the word $n+1$: This is known as the spillover effect (Rayner, 1998). However, it is worth noting that the lengthening of gaze durations when a detection occurs could also be explained by an alternative hypothesis: that gaze durations might be influenced by the planning of the response (Griffin, 2003; Morgan & Meyer, 2005). When the letter is detected, participants have to plan the motor movement to produce the detection response, which would translate into longer gaze durations when a detection occurred than when an omission occurred. Unfortunately, the present data do not allow us to confidently distinguish between these two hypotheses.

Although all of the previously mentioned results are in line with the predictions of the GO model, results for response latencies are not. According to the GO model, the time needed to identify the target letter is assumed to be independent of the word in which it is embedded (Healy, 1994; Moravcsik & Healy, 1995) and as outlined in the introduction, truncation of the distribution of letter processing times (due to interruption of search by the completion of word processing) will necessarily yield faster target detection times when the omission rate is higher. The results for response latencies are unequivocally contrary to this prediction. Response latencies were shorter for the content words than for *des* and for the subsamples of occurrences (1) of no regression; (2) of at least one regression; and (3) of skipped words. In addition, there was a positive correlation between the log frequency and the response latencies, indicating that the higher the frequency, the longer the response latency. There was also a positive correlation between omission rate and response latencies, indicating that the higher the omission rate, the longer the response latency. The longer response latencies for *des* than for its three-letter control content words cannot be attributed to a longer processing of *des*: When there was a detection, the gaze durations and total fixation times were the same for *des* and for the control words.

The present response latency results, together with similar previous results with the RSVP procedure (Saint-

Aubin & Klein, 2001; Saint-Aubin et al., 2003), highlight what might be called the Achilles' heel of the GO model: its prediction that target response latency will decrease as omission rate increases. This prediction derives from the following components of the unitization assumption: (1) independence—processing at the letter level is independent of processing at the lexical (word) level and (2) interruption—when processing at the lexical level is completed, processing of the word's constituent letters is interrupted as processing moves on to the next word. The combination of no influence of word difficulty on letter processing and interruption inexorably leads to the truncation of the letter-processing-time distribution at the point at which lexical processing is completed. The prediction that higher omission rates should be associated with faster target response latencies follows from the notion that there is one letter-processing time distribution that is truncated when processing at the lexical level is completed.

Saint-Aubin et al. (2003) noted that a model that maintains the interruption assumption of the GO model while dropping the independence assumption can accommodate an increase in response latencies alongside an increase in omission rate. This is illustrated in the upper right panel of Figure 1, where it is assumed that the processing of letters in frequent function words is less efficient (represented here as a delay in the response time distribution for possible target detections) than that of less frequent content words. It must be noted, however, that none of the extant models (which includes the original version of the GO model, the modification proposed by Saint-Aubin et al., 2003, and the structural account and unitization account/processing time hypothesis, on both of which the GO model is based) include an explicit mechanism for performing the search task. This is problematic, because the missing-letter effect is only observed when a letter search task is performed; and as mentioned above, there is a debate on the influence of the search task on the reading processes that can only be addressed by providing an explicit mechanism for the search task (Rayner & Pollatsek, 1989; Rayner & Raney, 1996). An additional problem encountered by interruption models like those illustrated in the upper panels of Figure 1 is that they make predictions about the distribution of detection response latencies that are not fulfilled. For example, because interruption entails truncation of the distribution, and because frequent function words have higher omission rates (hence, greater truncation), the within-subjects standard deviation of target detection times should be lower for frequent function words, whereas in fact the opposite result was obtained in the present study: Average within-subjects standard deviations were 265 msec and 237 msec for targets in function and content words, respectively, a difference (28 msec) in the wrong direction for interruption models that was significant [$t(40) = 2.05, p < .05$]. Similarly, under the interruption assumption, the response latency distributions should be negatively skewed, because the slow tail is truncated; however, our distributions, for both content and function words, are positively skewed. In the next section, we sketch out a computationally explicit model that

endorses the independence assumption while rejecting the interruption assumption of the GO model.

The Attentional-Disengagement Model

Our model is first and foremost a model of searching in the context of reading for comprehension. In this model, there is an explicit mechanism for performing the letter search task, which this mechanism has been designed to allow to timeshare with reading, for comprehension with minimal interference. As in many information processing models, it is assumed that information about the possible presence of a target accumulates while attention is engaged on the target containing a lexical item, and that this information begins to decay as soon as attention is disengaged. We have assumed that accumulation follows a power function, whereas decay follows an exponential function (see Figure 1E).³ Minimal interference with reading is accomplished by making this a passive process in which the probability of making a “target” response at a given point in time is directly proportional to the amount of information accumulated at that time.⁴ Once a response is made, the information accumulator is reset to zero. Monte Carlo simulation of this model yields patterns of target response latencies (Figure 1F) and omission rates that match those observed in the present studies.⁵

According to the attentional-disengagement model (AD), the proximal cause of the missing-letter effect is the timing of the disengagement of attention (Posner & Petersen, 1990) from the target-containing lexical unit. The use of the concept of attentional disengagement allows our model to link with a well established model of reading, the E-Z reader model, in which word recognition in reading is considered a serial process under the control of an attentional beam (Pollatsek, Reichle, & Rayner, 2003). All words would be attended to even if they were obviously not all fixated. Under this view, omissions would result from a rapid disengagement of attention from the target-containing word. Word function and word frequency would be distally related to the missing-letter effect by influencing the disengagement of attention. Because readers accurately predict the location of function words in a sentence, as suggested by the present results and those of Saint-Aubin et al. (2005), and because function words provide very little information about meaning (Koriat & Greenberg, 1994), readers would rapidly disengage their attention from them. It is also assumed that word frequency could influence the allocation of attention, since the disengagement of attention would be slower for low than for high frequency words. If the reader had to perform a letter search task while reading, it would be harder to find a target within a word from which attention is rapidly disengaged. This would translate into a higher omission rate and longer response latencies.

It is important to note the similarities and differences between the GO model and the AD model proposed here. As noted above, through inheritance of the assumptions of the unitization account-processing time hypothesis, the GO model assumes that letter and word processing are independent, and that processing at the letter level is interrupted in the reading process once the word has been

identified. As described in Saint-Aubin et al. (2003), this entails the truncation of the hypothetical target response latency distribution when omissions occur, thus predicting target response latencies to decrease as the omission rate increases. While maintaining the independence assumption of the GO model, the AD model does not assume that there is no chance of identifying a target letter once its lexical carrier is identified; instead, this model allows for the gradual decay (or *graceful degradation*, in neural network terms) of target information, once attention is disengaged from its lexical carrier. With the AD model, as with the GO model, factors that might cause a lexical item to be recognized more rapidly—bottom-up factors, such as perceptual fluency due to frequency, or top-down factors, such as expectancy or structure—will result in a greater omission rate because of the more rapid attentional disengagement.

Conclusion

Our results reveal that the pattern of eye movements in a reading task is not severely distorted by the presence of a letter search task, since benchmark effects found in the eye movement literature were reproduced here. This clearly establishes that the eye movement database can shed light on the missing-letter effect. In addition, by perfectly mirroring previous findings in the field with the RSVP procedure (Saint-Aubin & Klein, 2001; Saint-Aubin et al., 2003), the present results also reveal that, despite the fact the RSVP procedure finesses the oculomotor system which typically plays a role in reading, it is adequate for investigating the pattern of response latencies related to the missing-letter effect. The key finding is that the pattern of response latencies is contrary to the predictions of the GO model whether the text is presented in RSVP or more normally. We have presented an alternative model that predicts both errors and response latency data in the missing-letter effect. In future work, we plan to test the AD model by seeing how well it can predict a broader range of findings from the literature and the distribution of response latencies in new experiments with a considerable increase in the number of target trials.

AUTHOR NOTE

This research was supported by discovery grants from the Natural Sciences and Engineering Research Council of Canada to J.S.-A. and to R.M.K., and by a grant from the Canadian Language and Literacy Research Network (CLLRNet) to J.S.-A. and R.M.K. This article was part of A.R.-C.'s doctoral thesis. We thank Irina Holes for programming the experiment and Annie Jalbert and Isabel Savoie for their assistance in running participants and scoring the data. Correspondence concerning this article should be addressed to A. Roy-Charland or J. Saint-Aubin, École de Psychologie, Université de Moncton, Moncton, NB, E1A 3E9 Canada (e-mail: ear3339@umoncton.ca or saint-aj@umoncton.ca).

REFERENCES

- BALOTA, D. A., POLLATSEK, A., & RAYNER, K. (1985). The interaction of contextual constraints and parafoveal visual information in reading. *Cognitive Psychology*, *17*, 364-390.
- CARPENTER, P. A., & JUST, M. A. (1983). What your eyes do while your mind is reading. In K. Rayner (Ed.), *Eye movements in reading: Perceptual and language processes* (pp. 275-307). New York: Academic Press.

- CORCORAN, D. W. J. (1966). An acoustic factor in letter cancellation. *Nature*, **210**, 658.
- DRIEGHE, D., RAYNER, K., & POLLATSEK, A. (2005). Eye movements and word skipping during reading revisited. *Journal of Experimental Psychology: Human Perception & Performance*, **31**, 954-969.
- GREENBERG, S. N., HEALY, A. F., KORIAT, A., & KREINER, H. (2004). The GO model: A reconsideration of the role of structural units in guiding and organizing text on line. *Psychonomic Bulletin & Review*, **11**, 428-433.
- GREENBERG, S. N., KORIAT, A., & VELLUTINO, F. R. (1998). Age changes in the missing-letter effect reflect the reader's growing ability to extract the structure from text. *Journal of Experimental Child Psychology*, **69**, 175-198.
- GREENBERG, S. N., & TAI, J. (2001). Letter detection in very familiar texts. *Memory & Cognition*, **29**, 1088-1095.
- GRIFFIN, Z. M. (2003). A reversed word length effect in coordinating the preparation and articulation of words in speaking. *Psychonomic Bulletin & Review*, **10**, 603-609.
- HADLEY, J. A., & HEALY, A. F. (1991). When are reading units larger than the letter? Refinement of the unitization reading model. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **17**, 1062-1073.
- HEALY, A. F. (1976). Detection errors on the word *the*: Evidence for reading units larger than letters. *Journal of Experimental Psychology: Human Perception & Performance*, **2**, 235-242.
- HEALY, A. F. (1994). Letter detection: A window to unitization and other cognitive processes in reading text. *Psychonomic Bulletin & Review*, **1**, 333-344.
- KORIAT, A., & GREENBERG, S. N. (1991). Syntactic control of letter detection: Evidence from English and Hebrew nonwords. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **17**, 1035-1050.
- KORIAT, A., & GREENBERG, S. N. (1994). The extraction of phrase structure during reading: Evidence from letter detection errors. *Psychonomic Bulletin & Review*, **1**, 345-356.
- KORIAT, A., GREENBERG, S. N., & GOLDSHMID, Y. (1991). The missing-letter effect in Hebrew: Word frequency or word function? *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **17**, 66-80.
- MINKOFF, S. R. B., & RANEY, G. E. (2000). Letter-detection errors in the word *the*: Word frequency versus syntactic structure. *Scientific Studies of Reading*, **4**, 55-76.
- MORAVCSIK, J. E., & HEALY, A. F. (1995). Effect of meaning on letter detection. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **21**, 82-95.
- MORGAN, J. L., & MEYER, A. S. (2005). Processing of extrafoveal objects during multiple-object naming. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **31**, 428-442.
- NEW, B., PALLIER, C., FERRAND, L., & MATOS, R. (2001). Une base de données lexicales du français contemporain sur Internet: Lexique [A lexical database of contemporary French on the Internet: Lexique]. *L'Année Psychologique*, **101**, 447-462.
- O'REGAN, K. (1979). Saccade size control in reading: Evidence for the linguistic control hypothesis. *Perception & Psychophysics*, **25**, 501-509.
- POLLATSEK, A., REICHLER, E. D., & RAYNER, K. (2003). Modeling eye movements in reading: Extensions of the E-Z reader model. In J. Hyönä, R. Radach, & H. Deubel (Eds.), *The mind's eye: Cognitive and applied aspects of eye movement research* (pp. 361-390). Amsterdam: Elsevier.
- POSNER, M. I., & PETERSEN, S. E. (1990). The attention system of the human brain. *Annual Review of Neuroscience*, **13**, 25-42.
- RANEY, G. E., & RAYNER, K. (1995). Word frequency effects and eye movements during two readings of a text. *Canadian Journal of Experimental Psychology*, **49**, 151-173.
- RATCLIFF, R., & SMITH, P. L. (2004). A comparison of sequential sampling models for two-choice reaction time. *Psychological Review*, **111**, 333-367.
- RAYNER, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, **124**, 372-422.
- RAYNER, K., ASHBY, J., POLLATSEK, A., & REICHLER, E. D. (2004). The effects of frequency and predictability on eye fixations in reading: Implications for the E-Z reader model. *Journal of Experimental Psychology: Human Perception & Performance*, **30**, 720-732.
- RAYNER, K., & POLLATSEK, A. (1987). Eye movements in reading: A tutorial review. In M. Coltheart (Ed.), *Attention and performance XII: The psychology of reading* (pp. 327-362). Hillsdale, NJ: Erlbaum.
- RAYNER, K., & POLLATSEK, A. (1989). *The psychology of reading*. Englewood Cliffs, NJ: Prentice Hall.
- RAYNER, K., & RANEY, G. E. (1996). Eye movement control in reading and visual search: Effects of word frequency. *Psychonomic Bulletin & Review*, **3**, 245-248.
- REICHLER, E. D., POLLATSEK, A., FISHER, D. L., & RAYNER, K. (1998). Toward a model of eye movement control in reading. *Psychological Review*, **105**, 125-157.
- ROY-CHARLAND, A., & SAINT-AUBIN, J. (2006). The interaction of word frequency and word class: A test of the GO model's account of the missing-letter effect. *Quarterly Journal of Experimental Psychology*, **59**, 38-45.
- SAINT-AUBIN, J., & KLEIN, R. M. (2001). Influence of parafoveal processing on the missing-letter effect. *Journal of Experimental Psychology: Human Perception & Performance*, **27**, 318-334.
- SAINT-AUBIN, J., & KLEIN, R. M. (2004). One missing-letter effect: Two methods of assessment. *Canadian Journal of Experimental Psychology*, **58**, 61-66.
- SAINT-AUBIN, J., KLEIN, R. M., & LANDRY, T. (2005). Age changes in the missing-letter effect revisited. *Journal of Experimental Child Psychology*, **91**, 158-182.
- SAINT-AUBIN, J., KLEIN, R. M., & ROY-CHARLAND, A. (2003). Direct assessments of the processing time hypothesis for the missing-letter effect. *Journal of Experimental Psychology: Human Perception & Performance*, **29**, 1191-1210.
- SAINT-AUBIN, J., & POIRIER, M. (1997). The influence of word function in the missing-letter effect: Further evidence from French. *Memory & Cognition*, **25**, 666-676.
- SPRAGINS, A. B., LEFTON, L. A., & FISHER, D. F. (1976). Eye movements while reading and searching spatially transformed text: A developmental examination. *Memory & Cognition*, **4**, 36-42.

NOTES

1. More precisely, because with an RSVP procedure the presentation duration is fixed, even if its duration is selected to be similar to the usual fixation duration, for different participants or passages it might be either too long or too short. It has been suggested that the pattern of response latencies would not be a by-product of the cognitive processes highlighted by the missing-letter effect, but could be due to the discrepancy between the presentation duration and the time needed by the reader to process the word in view (see Healy's suggestion in Saint-Aubin & Klein, 2001, p. 332). According to this view, when presentation time is too fast, the reader could be exposed to a new word before the identification of the previous word is completed. Consequently, when the next word is displayed before the complete processing of the word which was in view, the reader could choose either to abort processing the word which was in view or to complete its processing. Because frequent function words are usually processed faster than content words, readers might choose more frequently to complete the processing of a function word, because its processing is advanced, before beginning processing the new word in view. This would produce longer response latencies for function words, because readers would choose to finish their processing (when the next word is displayed) more frequently than they would for content words.

2. For the norm collection, 549 participants took part in one of the five concurrent versions of the Web experiment for the predictability norms collection (see Saint-Aubin et al., 2005, for a detailed presentation of the Web experiment). Only 245 of them completed the entire experiment with a minimum of 44 for the fifth version and a maximum of 55 for the fourth version. Only the data from those who completed the entire experiment were used. Participants for the Web experiment were recruited by sending an e-mail message to all students at the Moncton campus of the Université de Moncton (Saint-Aubin et al., 2005). The message was sent to the e-mail address provided to the students by the university. After participants completed an online informed consent form and read instructions, the experiment began. Participants saw the first words of

the text and were allowed 20 sec to write their answer in an answer box. Once they gave their answer, the text was displayed along with their answer in green if it was the correct answer or in red if not. Then the text was presented with the next four words followed by the answer box and a new countdown was initiated. The online feedback provided to participants displaying the actual word was intended to mimic the normal reading situation in which the reader makes expectations about words, which are evaluated against the correct word once it is identified. For the norm collection, both the proportion of accurate prediction and the proportion of participants expecting a function word—the test word or not—were computed. The proportion of accurate prediction was computed for the 48 test words (24 *des* and 24 control content words) by dividing the number of times the test word was produced by the number of responses produced. Misspelled targets were considered correct as long as the test word was unambiguously recognizable. The proportion of participants expecting a function word was computed by dividing the number of times a function word was produced as a response by the number of responses produced. Following the list of function words presented by Koriat and Greenberg (1994), articles, prepositions, and conjunctions, as well as possessive and demonstrative adjectives, were considered function words (see Saint-Aubin et al., 2003, for classification in French of those adjectives as function words). Nouns, verbs, qualifying adjectives, adverbs, and personal pronouns were considered content words (Greenberg et al., 1998).

3. For the growth portion of the curve, a power function was chosen because its properties, including acceleration of the rate of accumulation, fit our initial expectations of what would be necessary, but a linear function works just as well. For the decay portion, an exponential decay function was chosen because this form characterizes decay in a multitude of natural contexts, including human memory. Again, subsequent investigations of a simpler linear decay have shown that the same general pattern of results can be obtained, though on a slightly less realistic scale.

4. This relationship may be thought to manifest biologically as an accumulator, represented by a single neuron or network of neurons, that gradually gathers input from lower processing levels and initiates a response when a certain threshold of input is reached. If this threshold were variable within trials, the relationship between information and response time would become probabilistic, as modeled here.

5. We have also shown that these results can be achieved by a homologous model using diffusion (see Ratcliff & Smith, 2004), in which there is noise in the accumulation process of a random walk toward a detection threshold and decay following disengagement.

(Manuscript received April 25, 2005;
revision accepted for publication August 14, 2006.)