Reading Long Words Embedded in Sentences: Informativeness of Word Halves Affects Eye Movements

Jukka Hyöna and Pekka Niemi
University of Turku, Turku, Finland

Geoffrey Underwood
University of Nottingham, Nottingham, England

The possibility was explored that the informativeness of a specific region within a word can influence eye movements during reading. In Experiment 1, words containing identifying information either toward the beginning or toward the end were displayed asymmetrically around the point of fixation so that the reader was initially presented with either the informative or noninformative zone. Words were read with shorter summed initial fixation time when the reading was started from the informative zone. In Experiments 2 and 3, the target words were presented in sentences that were to be comprehended. More attention was given to the informative endings of words than to redundant endings. The latter were also skipped more often. The duration of the first fixation was not affected by information distribution within the word, whereas the second fixation duration was. The results of these experiments lend good support to the hypothesis of immediate lexical control over fixation behavior and to the notion of a convenient viewing position.

Three experiments deal with two related issues in the investigation of eye guidance during reading. First, does the information distribution within a word influence fixation time on that word? Second, are fixation locations affected by the information distribution within words? These questions were investigated by observing the eye fixations of readers who were presented with words containing uneven distributions of information. The analyses are based on fixations in two different locations: those coinciding with the informative part of a target word and those coinciding with the uninformative part.

There is evidence that the information value of a word may affect the pattern of eye movements during reading. It has been demonstrated that syntactic constraints can affect fixation patterns. Carpenter and Just (1983) and O'Regan (1979) reported that subjects fixated three-letter function words less frequently than content words of equal length. An effect of semantic constraint was observed by Ehrlich and Rayner (1981). Specific words in predictable contexts were fixated less often than the same words in neutral contexts, and when they were fixated, they received shorter fixations. These are examples of how words that carry more information are fixated more often and longer than more redundant words.

It is also possible that different parts of a single word may carry different amounts of information. Could this be reflected in readers' fixation patterns? This question was first posed by O'Regan (1984), who used target words long enough to attract more than one fixation. The distinctive feature of these words was that information vital for identifying the word was located either within the first six letters or within the final six letters. The readers were asked to judge whether two words were semantically related. The target words were presented asymmetrically about the point of fixation, with fixation coinciding either with the third letter from the beginning or the third letter from the end of the word. The idea is that if fixations are under the moment-to-moment control of ongoing processing, then being forced to fixate a noninformative part of a word makes the reader move quickly away from that part. Therefore, the informative part of the same word would receive longer fixations. The findings of O'Regan and his associates lent partial support to this hypothesis. The information-at-the-end [info(end)] words were read with relatively short gaze durations when reading started from the end (i.e., from the informative zone). Gaze durations were slightly longer when reading started from the beginning of the word. Analogously, the information-at-the-beginning [info(beg)] words received short gaze durations when reading started from the beginning of the word, and longer gaze durations when the initial point of fixation coincided with the less informative end zone.

O'Regan (1984) performed a separate analysis concerning eye movement patterns involving either one or two fixations on a word. When reading started from the beginning, 22% of the info(beg) words were read with a single fixation, whereas...
only 9% of the info(end) words displayed this pattern. This means that the redundant, noninformative part of the word is sometimes skipped, but the informative part is skipped less often.

The results were less consistent when the end parts of the words were fixated. In fact, an initial fixation at the end of a word was always short and was followed by a longer fixation at the beginning, regardless of the information content of the end part. Thus O'Regan et al. (1984) confirmed the idea of "left dominance" of the initial fixation location. It was concluded that the convenient viewing position was between the third and fifth letter. The convenient viewing position, first described by O'Regan (1981), refers to the fact that a word is read with shortest gaze durations when the initial fixation is forced slightly left from the word's center. In normal, continuous reading, the eyes tend to land initially near the word's center (see Dunn-Rankin, 1978; Kliegl, Olson, & Davidson, 1983). This phenomenon is referred to as the preferred viewing position, first described by Rayner (1979). In other words, the convenient viewing position is the optimal position for beginning to inspect a word, whereas the preferred viewing position is the actual landing position on a word.

The ecological validity of the findings by O'Regan (1984) and O'Regan et al. (1984) can be questioned. However, it is possible to develop their paradigm further so as to better tap the process of natural reading. If the words are embedded in real sentences, no question arises about the role played by the forced nature of the initial fixation. More specifically, when a sentence is being read, the eyes are free to begin inspecting any part of the critical word. In the following study, Experiment 1 aims to replicate the semantic comparison results of O'Regan and his associates. In Experiments 2 and 3, the target words are embedded in sentences that comprise a brief story. Two specific predictions are made for Experiments 2 and 3. First, if fixations are under the moment-to-moment control of on-going information processing, then the informative parts of the target words should receive more visual attention than the redundant parts. Second, the noninformative word parts should be skipped—that is, left unfixedated—more often than the informative word parts.

Experiment 1: Reading Isolated Words

Method

Stimuli. Eighty Finnish words were selected for a pilot study conducted in order to find words identifiable either from the initial or the final six letters. In this preliminary experiment, word length was between 10 and 13 characters. In 40 of the words, the information assumed to be crucial for identification was word-initial; in the other half this information was located at the end. Two 80-word lists were constructed. A given word appeared in both. In one list, the six initial letters of the word were visible; the other list displayed only the six final letters. Both types of presentation (initial six letters visible, final six letters visible) appeared in each list equally often. Subjects were instructed to complete each word. They were informed that each word contained at least 10 letters, but no indication was given of the exact length of any word. Twenty subjects performed the completion task with one list, and another 21 subjects were given the other list.

The final battery consisted of 20 words with crucial information at the beginning and 20 words with crucial information at the end, chosen on the basis of this preliminary experiment. The selection procedure worked satisfactorily. For info(beg) words, the rate of correct identification was 96.6% when only the first six letters were shown. With the final six letters shown, the identification dropped to 9.0%. The respective identification rates for info(end) words were 6.1% and 86.4%.

When a word is here described as containing its crucial information zone at the beginning, that means it can be identified from its first six letters but not from its last six letters. Similarly, when a word is described as containing its crucial information zone at the end, it can be identified from the last six letters but not from the first six. The final set of words used in Experiment 1 is given in Appendix A.

In addition to the informativeness of the zones, there is a second within-word factor that is relevant here. If words include orthographically highly redundant sequences, this may well cause the reader to skip over such segments or allocate just a short fixation to them. To control for this possibility, a bigram frequency count was performed for both halves of each target word. The reference is 24,000 bigram occurrences in three different types of Finnish prose (Mikkonen, 1972). The results are given in Table 1. The data consist of means across word halves in a given word category, and the means were based on an average bigram frequency within an individual word half.

The most redundant part was the latter half of those words which had the information at the beginning. It differed statistically from the other three word halves, among which there were no significant differences.

The mean length of info(beg) words was 11.45, and that of info(end) words was 10.60 characters. All info(end) words were compound words, whereas this was the case for 12 of the info(beg) words. Each compound word comprised two nouns. This type of compound word occurs very frequently in Finnish. The noncompound words [i.e., 8 info(beg) words] were adjectives. All the 20 info(beg) words contained a frequent suffix at the end.

To control for the familiarity of the stimulus words, a familiarity rating was performed by using a method similar to Gernsbacher's (1984). A separate group of 21 subjects, chosen from the same population as the experimental groups, judged the familiarity of all the 40 target words on a 7-point scale, in which the figure 1 denoted a very familiar word and the figure 7 a very unfamiliar word. This familiarity judgment task revealed that there was no difference in the experienced familiarity between the info(beg) and info(end) words. The average familiarity for both info(beg) and info(end) words was 3.5 on a 7-point scale [SD = 1.53 for info(beg) and SD = 1.29 for info(end) words, respectively].

Apparatus. The words were displayed on a conventional monochrome TV screen (Helvar Apollo TV-53) and were seen as white against a dark background. The displayed words were in uppercase letters. The width of one character was 0.93 cm; with a viewing

<table>
<thead>
<tr>
<th>Target word</th>
<th>Beginning IL</th>
<th>End IL</th>
</tr>
</thead>
<tbody>
<tr>
<td>First half</td>
<td>104.9</td>
<td>126.8</td>
</tr>
<tr>
<td>Second half</td>
<td>227.3</td>
<td>131.4</td>
</tr>
</tbody>
</table>

Note: The larger the value, the more common the bigram combination of Finnish.
distance of 62 cm, one character subtended a visual angle of about 55 min of arc. Eye fixations were collected by an Applied Science Laboratories (ASL) Eye-Trac Model 200, which is a spectacle-mounted photoelectric recorder. The measurement accuracy of the apparatus, as reported by the manufacturer, is 1° horizontally and 2° vertically. Eye positions were sampled every 10 ms by an Apple Ile microcomputer. A chin-rest and a headband were used to restrict head movements.

Procedure. Prior to the presentation of the battery of words, each subject took part in a calibration procedure by using the Eye-Trac apparatus and three squares displayed on the TV screen. These squares delineated the maximum horizontal coordinates of the words used in the experiment, and each square was the size of one character space. The center square also served as the fixation point in the actual experiment and was midway between the other two squares. The subjects were asked to fixate the squares in sequence so that the experimenter could perform an accurate calibration of the Eye-Trac apparatus.

A trial in the experiment began by the appearance of the fixation square in the center of the TV screen. Its location was the same as the third character from the beginning of a word (initial fixation at word beginning) or the same as the third character from the end of a word (initial fixation at word ending). When the subject had been fixating the square for about 100 ms, it was replaced by the target word. The word that appeared was displayed asymmetrically around the fixation point, with the third letter from the beginning or the third letter from the end coinciding with the fixation point. The word remained visible until the subject pressed a button in order to gain sight of a comparison word that appeared eight character spaces below the first word. The subject’s task was to indicate whether the two words were synonyms. The response was given by pressing a yes or a no button. After an interval of 3 s, the fixation square reappeared to start the next trial. If minor departures of calibration occurred (indicated by nonfixation of the fixation square), the experimenter readjusted the recorder while the subject looked at the square.

Data were collected only from the first word presented on each trial. The synonymity judgment was employed to ensure that the subjects read the words for meaning. All 40 trials were presented in one block, with order of presentation randomized. Half of the subjects read a given word with the initial fixation towards its beginning, and the other half read the same word with the initial fixation towards its ending. Prior to the experimental trials, subjects were presented with 10 practice trials, each consisting of a word pair discarded from the pilot study.

Design. The experiment employed a 2 x 2 within-subjects repeated measurements design. The variables were information location (beginning/end of a word) and initial fixation location (beginning/end of a word). Ten words were presented in each of these four conditions.

Subjects. Twelve Finnish-speaking adults (9 females) served as subjects for the experiment. All had normal vision. The data from 1 subject had to be discarded from the analysis because of equipment failure.

Results

Preliminary observations indicated that microsaccades and drifts smaller than one character space occurred in the course of the experiment. This happened in particular when the subject was viewing the fixation square. If such saccades are taken as a borderline between two consecutive fixations, the procedure would result in short average fixations, the psychological status of which could be questioned. Fixations with a distance of less than one character space were therefore treated as a single fixation. Fixations shorter than 100 ms were discarded from the analysis.

**Summed initial fixation time.** The first analysis concerned the summed initial fixation time. The occurrence of a regression terminated the summation, except when the subject was forced to start the reading from the end of the word. The conventional gaze duration measure was not applied because of the experimental procedure, which encouraged subjects to move back and forth on the target word while preparing to press the button.

An analysis of variance (ANOVA) was computed on the summed initial fixation time data, with two within-subjects factors: information location and initial fixation location. The ANOVA indicated a main effect of initial fixation location, *F*(1, 10) = 9.05, *p* = .013, *MS* = 10,668. The summed fixation time was longer when reading started from the end of the word. There was no main effect associated with information location, but it was involved in an interaction with initial fixation location, *F*(1, 10) = 6.54, *p* = .029, *MS* = 5,757. With the initial fixation at word beginning, the summed fixation time was 733 ms for info(beg) words and 755 ms for info(end) words. When the initial fixation was at word ending, the summed fixation time was 885 ms for info(beg) words and 720 ms for info(end) words.

**Fixation strategies.** Words gained varying numbers of fixations, and these data are presented in Table 2. The data consist of probabilities of reading a target word with one, two, three or more fixations. These data were subjected to an ANOVA involving three within-subjects factors: fixation frequency, information location, and initial fixation location. There was a main effect of fixation frequency, *F*(3, 30) = 19.94, *p* < .001, *MS* = 0.08, suggesting that the two-fixation strategy occurred more frequently than the others. More interesting is the interaction between fixation frequency and initial fixation location, *F*(3, 30) = 14.85, *p* < .001, *MS* = 0.05. When reading commences with a fixation on the initial part, a word is most often read with two fixations. When reading starts from the end of the word, the reader usually needs three fixations.

The durations of the first two fixations were subjected to a separate ANOVA involving three factors: information location, initial fixation location, and first/second fixation upon a word. The results are given in Figure 1. Only those effects are reported in which the factor of first/second fixation was

<table>
<thead>
<tr>
<th>Table 2</th>
<th>Probabilities of Fixation Distributions in Experiment 1, as a Function of Information Location and Initial Fixation Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fixation frequency</td>
<td>Info(beg)</td>
</tr>
<tr>
<td>Fix (beg)</td>
<td>0.04</td>
</tr>
<tr>
<td>Fix (end)</td>
<td>0.62</td>
</tr>
<tr>
<td>3</td>
<td>0.29</td>
</tr>
<tr>
<td>≥ 3</td>
<td>0.05</td>
</tr>
</tbody>
</table>

Note. Info = information; Fix = fixation; Beg = beginning.
involved, because the summed duration of the two fixations is uninteresting in the absence of data on the possible third and fourth fixations.

First/second fixation failed to produce a main effect, but it was involved in two interactions. First/second fixation x initial fixation location, $F(1, 10) = 16.15, p = .002, MS_e = 53,137$, indicates that when reading was started from the end part of the word, a very short first fixation occurred. This interaction suggests that the subjects have a preference to move quickly to the initial part when they are forced to start reading the word from the end.

More interesting is the interaction between first/second fixation, initial fixation location, and information location, $F(1, 10) = 11.44, p = .007, MS_e = 15,813$. Post hoc Scheffé tests showed that when reading started from the beginning of the word, the durations of the first and second fixation did not display any reliable difference associated with information location. A different picture emerged when reading started from the end of the word. For info(beg) words, the second fixation was significantly longer than the first ($p < .05$) because it was directed at the informative zone. For info(end) words this was not the case, and a shorter second fixation resulted.

**Initial saccade length.** An additional dependent measure recorded was the length of the first saccade, measured in character spaces from the fixation point. The analysis was computed on the absolute values of saccade length. When the initial fixation was on the word beginning, some rare trials resulted in a leftward first saccade (less than 1%); these were discarded from the analysis. Analogously, the rare rightward first saccades after the initial fixation on the word ending (less than 0.5%), were also discarded from the analysis. When the initial fixation was at word beginning, the length of the first saccade was 3.2 character spaces for info(beg) words and 3.5 for info(end) words. With the initial fixation at word ending, the length of the first saccade was 5.2 character spaces for info(beg) words and 4.3 for info(end) words. The effect of initial fixation location was again highly significant, $F(1, 10) = 78.81, p < .001, MS_e = 92.6$. Longer saccades resulted from the initial fixation at the end of a word. Information location also produced a significant main effect, $F(1, 10) = 8.30, p = .016, MS_e = 40.8$. On the average, the first saccade was longer for words whose informative zone was at the beginning. However, this main effect is qualified by the occurrence of a

**Discussion**

In Experiment 1, the subject inspected one word that was presented asymmetrically around the point of fixation, in preparation to performing a synonymity judgment task. The data on summed fixation duration quite closely replicated the results reported by O'Regan et al. (1984), with gaze duration as the dependent measure. Because of the differences in experimental procedure, the “gazes” in the present experiment were, in general, more than 200 ms longer. In accordance with O'Regan et al. (1984), we found the beginning dominance phenomenon, which refers to the simple fact that starting the reading of a word from the end requires extra effort. The same phenomenon was also revealed by the differences in fixation patterns. Backward reading resulted in more fixations than forward reading. What is even more interesting, we found, as did O'Regan et al. (1984), that the effect was much larger for info(beg) words than for info(end) words. This means that by fixating first on a redundant word ending, the subject was able to extract very little information about the identity of the word. On the other hand, the subject seemed to grasp some of the word’s identity by initially fixating the informative ending of the word.

The target words were almost never read with a single fixation, which runs counter to O'Regan's (1984) observations. This can be explained by the differences in the experimental procedure. The comparison word appearing outside the parafoveal region after a button press discouraged our subjects to move immediately away from the target. Starting the reading from the word ending resulted in more fixations, but there was no evidence that the location of information affected the fixation pattern. This seems to suggest that informative zones were not fixated more frequently or noninformative zones more seldom. In fact, this could have been expected on the basis of O'Regan's (1984) findings.

The first fixation duration proved not sensitive to the location of the informative zone within a word (see Figure 1, left panel). The longest first fixations were recorded when the eye initially fixated the beginning of a word, regardless of the location of information. The second fixation, however, did show some sensitivity to the location of the informative zone. Recently, O'Regan and Lévy-Schoen (1987, Figure 7) have reported a similar finding. When the initial fixation fell on a noninformative zone, the second fixation was always long (see Figure 1, right panel). This may represent an escape from a noninformative zone, but it does not explain why the initial fixation at the beginning of the word was long relative to the initial fixation at the end. However, the data do eliminate the possibility that an imbalance in letter bigram frequencies was operative (see Table 1). The initial fixation at the end of a word was of equal length, regardless of whether the end hosted an informative zone or a redundant, noninformative zone.
As noted earlier, the length of the first saccade also seems sensitive to the location of the informative zone within a word. The first saccade was longer when the initial fixation fell on the end of a word, which may reflect an attempt to recover the early letters of the word. Most important, when the locus of the initial fixation was at the end of the word, a longer saccade was made if the crucial information was at the beginning of the word. O’Regan (1984) also reported evidence that the length of the first saccade is affected by the location of information. However, his observation was restricted to target words read from left to right.

**Experiment 2: Reading Words Embedded in a Sentence**

Experiment 2 was planned to avoid some of the shortcomings of the paradigm used in Experiment 1. Starting reading from a predetermined location other than the word beginning is admittedly artificial. It also adds to the complexity of the design and results. In Experiment 2, the info(beg) and info(end) words used in Experiment 1 were embedded in sentences comprising brief stories (see Appendix B). Readers were thus allowed to approach these words more naturally.

**Method**

**Stimuli and apparatus.** The target words used in Experiment 2 were the same as in Experiment 1, but each word was incorporated in a sentence (4–9 words). Some of the words were modified by the inflectional changes required by Finnish syntax. The inflections reflect the case role the words occupied in the sentence. Twelve info(beg) and 7 info(end) words lacked any inflections. When inflected, 5 info(end) and 4 info(beg) words were marked as the sentence object (the inflection -a or -n), and 3 info(end) and 1 info(beg) word were in the genitive form (the inflection -n). The remaining 8 inflections were scattered over multiple categories.

Word length in Experiment 2 ranged from 10 to 16 letters. The mean length of info(beg) words was 12.00, and that of info(end) words was 11.85 letters. The target word never occupied the initial position in a sentence. The materials were presented in the form of brief stories, with three or four sentences per story. The words were displayed in uppercase letters. The presentation of sentences on the TV screen required a greater viewing distance than in Experiment 1, and with a distance of 89 cm, one character subtended a visual angle of about 37 min of arc. The apparatus was the same as in Experiment 1.

**Procedure.** Three squares were again used for the initial calibration: one square in the place of the first character of a line, one in the center, and one at the end of a line. A story trial began when the subject pressed a button. This caused a fixation square to appear at the beginning of a line. After the subject had fixated this square for about 100 ms, it was replaced with the first line of the story. After reading it, the subject pressed the button to obtain the fixation square, and then the next line of the story appeared in the same location. When the presentation of the story had been completed, the subject performed a comprehension task. This required a yes or no answer to a question dealing with the general topic of the story. The comprehension task was followed by recalibration. Prior to the experiment, each subject was presented with three practice stories. Each subject was subjected to an ANOVA, these data produced a main effect,

**Results**

**Gaze duration.** The gaze duration for a target word was defined as the total fixation time upon the word before leaving it. As in Experiment 1, the minimum fixation time accepted was 100 ms. In the ANOVA computed on the gaze duration data, information location produced a main effect, $F(1, 11) = 5.02$, $p = .047$, $M_S = 87,926$. The average gaze duration for info(beg) words was 425 ms, and that for info(end) words was 486 ms.

**Fixation strategies.** The probability of reading the target word with one, two, or more fixations did not show any effect related to information location, $F(1, 11) < 1$. As in Experiment 1, there was only the main effect of the fixation frequency, $F(3, 33) = 30.99$, $p < .001$, $M_S = 0.03$. The probability of fixating the target word only once was 0.30; twice, 0.50; three times, 0.14; and more than three times, 0.06. In other words, the most typical pattern consisted of two fixations. A separate ANOVA using subject means as the dependent measure was performed on the fixation durations resulting from these trials. The second fixation was marginally longer than the first fixation (242 ms vs. 202 ms), $F(1, 11) = 4.44$, $p = .06$, $M_S = 4,061$. The duration of the second fixation also showed some sensitivity to the information location, its duration being 260 ms for info(end) words and 224 ms for info(beg) words. However, the interaction was nonsignificant, $F(1, 11) = 2.02$. The duration of the first fixation was 210 ms for info(end) words and 193 ms for info(beg) words. Information location produced a significant main effect on the gaze duration associated with a two-fixation reading pattern, $F(1, 11) = 8.06$, $p = .018$, $M_S = 1,885$. The average gaze duration for info(beg) words was 418 ms, and that for info(end) words was 470 ms.

**Inspection time.** The next analysis dealt with the question of how fixations were distributed between the two halves of the target word. The inspection time for a word half was defined as including all individual fixations upon it before leaving the whole word. The data are presented in Table 3.

In the ANOVA computed on the inspection times, only the interaction between information location and word half reached significance, $F(1, 11) = 5.49$, $p = .039$, $M_S = 43,559$. The two word categories showed a large difference of inspection time upon the second half of the target words.

**Probability of fixating a word-half.** The probabilities of fixating each word half are presented in Table 4. When subjected to an ANOVA, these data produced a main effect of word half, $F(1, 11) = 24.24$, $p < .001$, $M_S = 0.05$. The first half of a word was more likely to be fixated than the second half. No effect of information location was observed, $F(1, 11) < 1$, and there was no significant interaction between the two factors, $F(1, 11) = 2.29$, although there was a tendency to skip the noninformative end part more often.

**First fixation duration.** A further ANOVA indicated that the duration of the first fixation upon the target word was not significantly different for info(beg) and info(end) words, $F(1, 11) = 1.52$. The respective means were 212 ms and 225 ms.
In accordance with Experiment 1, the average duration of the first fixation was not sensitive to the distribution of information. This was also true of the two-fixation strategy. However, the two-fixation strategy produced second fixations showing some sensitivity to the location of information, as in Experiment 1.

The number of fixations upon a target word was not affected by the distribution of information within that word. This runs counter to the finding by O’Regan (1984) that info(beg) words were read with one fixation more frequently than info(end) words. The main difference between the experiments was that we employed a natural reading paradigm. This resulted in 30% of the target words read with one fixation as opposed to 16% in O’Regan (1984).

Contrary to our hypothesis, the probabilities of fixating informative versus noninformative zones did not differ significantly from each other. In particular, we expected that redundant endings would have been left unfixed more often than informative endings. Although such a trend was visible in the data, it was not reliable. A possible explanation for the absence of the effect is that 12 (8 info[end] words) out of 40 target words were located at the end of a line. During the reading of such a word, rightward saccade programming is discontinued at some point, which may result in a pronounced beginning dominance and relatively few fixations at the end part of the word. Indeed, the probability of fixating the second half was 0.56 for the 12 target words and 0.63 for those located in the middle of a line.

Another possible explanation is that although info(beg) words sometimes are encoded on the first fixation, there may be a general tendency for long saccades to undershoot. Thus many of them would actually be intended for the next word but fall short and hit upon the word ending. This would mean longer initial saccades within a word for info(beg) words than info(end) words. However, no such trend was observed in the data.

The preferred landing position for long words seems to be slightly left from the word’s center, the fourth or fifth letter as found in Experiment 2. This corresponds quite closely to the convenient viewing position, that is, the most optimal landing position, as determined by O’Regan et al. (1984). Consequently, it appears that the distinction between the convenient and preferred viewing position is conceptual rather than empirical, albeit O’Regan and Lévy-Schoen (1987) have some evidence that they could be distinguished also empirically (see their Figure 16.10, on p. 379).

The finding that the initial fixation was located further toward the end of the word when the informative zone was at the end, is rather provocative. If reliable, it would suggest that informative or interesting locations could attract eye fixations. This attraction hypothesis, suggested by Underwood (1981, 1985) entails an assumption of semantic preprocessing of parafoveal words. The only reason why informative zones should attract eye fixations is the assumption that the eye guidance system has knowledge of the informativeness of these zones prior to their fixation. For instance, the end parts of info(beg) words could be recognized parafoveally as redundant word endings, which would thus result in the initial fixation to be guided more toward the word’s beginning.
However, the observed effect was small (0.65 character spaces) in relation to the accuracy of the recording apparatus. Consequently, a replication is needed before any firm conclusions can be drawn. It is also worth noticing that the whole issue of semantic preprocessing has been challenged (see Rayner, Balota, & Pollatsek, 1986).

Experiment 3: Reading Words Embedded in a Sentence

Experiment 3 was planned to replicate the results of Experiment 2, and differed from it in three respects. First, in Experiment 3, a target word was never placed as the last word in a sentence or a line, which was the case for some of the target words in Experiment 2. This was meant to prevent the process of “wrapping up” sentences from taking place during the reading of the target words. Just and Carpenter (1980) have suggested that long fixations tend to occur at the ends of sentences while readers compute the full meaning of the sentence. Second, info(beg) and info(end) targets were matched for word length. Minor changes were made in word order, wording, or syntax to meet these requirements. Third, this time lowercase letters were used. Otherwise, the stories used were the same as in Experiment 2.

Method

Stimuli. The target words were the same as in the previous experiments. In Experiment 3, each word was incorporated in a sentence of 4 to 13 words. As in the materials of Experiment 2, sentence structure necessitated inflectional changes in some of the words. Eleven info(beg) and 4 info(end) words were uninfl ected. Six info(end) words had an inflection of -a or -n, which denotes that the word is the sentence object. The remaining inflections were scattered over a variety of categories.

The length of the target words ranged from 10 to 16 characters, with a mean of 12.25 for both word categories. The target word never occupied the initial or the final position in a sentence or a line. The materials were presented in the form of brief stories, with three or four sentences per story. Lowercase letters were used. Each story was divided into pages comprising two lines of text.

Apparatus. The stories were displayed on a conventional TV screen (3inlux 510) and were seen as white against a dark background. The width of one character was 0.78 cm; with a viewing distance of 89 cm, one character subtended a visual angle of 30 min of arc. The text lines on each page were 2.5 cm apart, which equals a vertical visual angle of 1.6°. Eye movements were collected by an Applied Science Laboratories Model 1994. This monitoring system is video-based, and it tracks the pupil for horizontal eye movements and corneal reflection for vertical eye movements. The accuracy of the apparatus is 1° as reported by the manufacturer. A chinrest and a headband were used to restrict head movements.

Procedure. The initial calibration was identical to that in Experiment 2, except for one modification. Instead of one text line, two lines of text were presented at a time. Therefore, six calibration squares were needed. A story trial began when the subject pressed a button to obtain the two initial lines of text. When the subject pressed the button again, another two lines of text appeared. When the presentation of the story had been completed, the subject performed a comprehension task. This required a yes or no answer to a question dealing with the general topic of the story. The comprehension task was followed by a calibration check. Prior to the experiment, three practice stories were presented. Each subject was shown the same material.

Design. The experiment used a within-subjects repeated measures design. The factor was information location (beginning/end of the word), and 20 words were employed in each of the two conditions.

Subjects. Fifteen Finnish-speaking subjects (13 females) were selected from the same population as in the previous experiments. None had participated either in Experiment 1 or Experiment 2.

Results

Gaze duration. The gaze duration data from Experiment 3 indicated that info(beg) words were read with significantly shorter gaze than info(end) words, F(1, 14) = 5.97, p = .028, MS* = 38.749. The mean gaze duration was 457 ms for info(beg) words and 496 ms for info(end) words. Experiment 3 thus replicated the finding from Experiment 2.

Fixation strategies. The probability of reading the target word with one, two, three, or more fixations showed only the main effect of fixation frequency, F(3, 42) = 27.07, p < .001, MS* = 0.04. The probability of fixating the target word only once was 0.19; twice, 0.52; three times, 0.20; and more than three times, 0.09. There was no main effect of information location, nor was it involved in any interaction. These data are a close replication of results obtained in Experiment 2. The most typical fixation pattern, that is, reading the word with two fixations, was again subject to a separate analysis using subject means. Unlike in Experiment 2, the factor of information location did not produce a significant effect on gaze duration, which was 427 ms for info(beg) words and 431 ms for info(end) words. However, the durations of the first and second fixation showed an interaction with information location, F(1, 14) = 5.05, p = .041, MS* = 838. For info(beg) words, the duration of the first fixation was 228 ms, and that of the second fixation was 200 ms. The respective means for info(end) words were 212 ms and 218 ms.

These data suggest that the duration of the first and second fixation both would be affected by information location, whereas the data of Experiments 1 and 2 suggest that only the duration of the second fixation would be sensitive to information distribution. Because of this inconsistency, the two-fixation strategy was examined further by pooling the data from Experiments 2 and 3 together. The pooling revealed a reliable interaction between the duration of the first/second fixation and information location, F(1, 25) = 7.18, p = .013, MS* = 4,888. Although the duration of the first fixation was not sensitive to information location (213 ms for info(beg) and 211 ms for info(end) words), the duration of the second fixation seemed to be affected by it (210 ms for info(beg) and 236 for info(end) words). The second fixation, which was generally located on the latter word half, was longer when it was upon the informative zone. There was also a main effect for information location, F(1, 25) = 4.09, p = .054, MS* = 3,756, suggesting that info(beg) words were read with shorter gazes. This is due to the difference in the second fixation duration.

Inspection time. An analysis similar to that in Experiment 2 was performed on word halves. The inspection time data are presented in Table 3. All fixations prior to leaving the
whole word were included. The main effect of information location proved significant, $F(1, 14) = 6.48, p = .023, MS_E = 1.426$. In other words, info(end) words were slower to read than info(beg) words. The interaction between information location and word half was significant, $F(1, 14) = 8.26, p = .012, MS_E = 1.872$. The inspection times for info(beg) and info(end) words differed from each other only within the second half of the word, which replicated the finding of Experiment 2.

There is a possibility that the aforementioned effect on inspection time is due to the difference in the syntactic form class between info(beg) and info(end) words. Although all the info(end) words were compound nouns, 8 info(beg) words were adjectives, and the other 12 were compound nouns. To rule out this possibility, the inspection time data for 8 info(beg) adjectives were compared with the rest of info(beg) words. An ANOVA was performed by using subject means and by pooling the data of Experiments 2 and 3. The analysis showed that the form class did not interact with word half ($F < 1$). That is, there was no indication that the two word types would have affected the inspection times differently. For both word types the inspection time was longer on the first half, that is, on the informative zone. For adjectives the inspection time on the first half was 273 ms and on the second half, 242 ms; for compound words the respective means were 297 ms and 245 ms.

**Probability of fixating a word-half.** The probabilities of fixating a word half are given in Table 4. A significant main effect of word half emerged, $F(1, 14) = 5.65, p = .032, MS_E = 0.03$. Again, the result was due to the beginning dominance. Unlike in Experiment 2, this time the interaction between information location and word half was also significant, $F(1, 14) = 15.21, p = .002, MS_E = 0.005$. It was due to the particularly strong beginning dominance for info(beg) words.

**First fixation duration.** The duration of the first fixation was longer for info(beg) words (225 ms) than for info(end) words (211 ms). Surprisingly, this small difference proved significant, $F(1, 14) = 8.21, p = .013, MS_E = 3.732$.

**Initial saccade length.** There was no reliable difference between the two word categories in terms of the length of the first saccade within a word, $F(1, 14) < 1$; thus the finding of Experiment 2 was replicated. The mean algebraic saccade lengths were 3.8 character spaces for info(beg) words and 4.0 for info(end) words.

**Preferred landing position.** The position of the first fixation upon the target word showed only a marginal main effect of information location, $F(1, 14) = 3.17, p = .097$. The average positions to the left from the word center were 2.4 character spaces for info(beg) words and 2.1 for info(end) words.

**Discussion**

In general, Experiment 3 replicated the findings of Experiment 2. The inspection time data were almost identical. As in Experiment 2, both inspection times on word halves and gaze durations were affected by information location within a word. Hence, the possible confounding due to the sentence wrap-up phenomenon can be ruled out. In other words, the data of Experiment 2 were not affected by the fact that 12 out of 40 target words were located at the end of a sentence. Nor was the syntactic form class responsible for the observed results.

There is one more alternative interpretation for the inspection time data that has to do with bigram frequencies within a word. If bigram frequency were the source of the variation in inspection times, then, other things being equal, this would mean fast reading of the end part of info(beg) words in our Experiments 2 and 3, whereas the other three word halves should be read at a mutually equal but slower speed. However, our data are at odds with this notion.

In Experiment 3, the average duration of the first fixation seemed sensitive to the location of information. Although reliable, the effect was rather small (14 ms). In fact, together with the data of Experiments 1 and 2 and the pooled data for the two-fixation strategy, we have to conclude that there is little evidence supporting the notion that the duration of the first fixation would be affected by the distribution of information within a word. However, we do have consistent evidence that the duration of the second fixation displays sensitivity to information distribution. This is compatible with O'Regan and Lévy-Schoen (1987; see Figure 7), who found that when a word was read with more than one fixation, the duration of the first fixation was insensitive to information distribution within words, whereas the duration of the second fixation was affected by it.

O'Regan and Lévy-Schoen (1987) propose that the first fixation duration might be sensitive to lexical processing only in those cases when it is the only fixation on a word. Indeed, Lima and Inhoff (1985) reported data showing that the duration of the first fixation can be affected by lexical information, in their case trigram frequencies. They used five-letter words, which are often read with a single fixation. In the present study, however, this notion cannot be tested because of an insufficient amount of data points concerning the single-fixation strategy.

The data of Experiment 3 differed from those of Experiment 2 in one interesting aspect. Probability of fixating the first versus the second word half was not affected by the information location in Experiment 2. However, this was clearly the case in Experiment 3. Beginning dominance was virtually absent for info(end) words. That is, the informative end zones of the words were regularly fixated, whereas the redundant endings were sometimes skipped over. Also the probability of fixating the end part of a word was higher in Experiment 3. It appears that the change was due to the fact that no target words were located at the end of a line in Experiment 3. This time the evidence for the aforementioned attraction hypothesis—that is, the influence of information distribution on the preferred landing position—failed to prove reliable, although the effect was in the same direction as in Experiment 2.

**Conclusions**

A number of factors may influence the duration and location of an eye fixation during reading, and the experiments reported here have identified the importance of the location
of information within a word and a preferred landing position within a word. Factors not considered here include, for example, syntactic and semantic constraints; thus, our conclusions are based upon a restricted view of the process of eye guidance during reading.

In Experiment 1 the first fixation was remarkably brief if it was located at the end of the word. The information content of the word half had no effect on its duration, and hence this finding is similar to the "beginning dominance" phenomenon reported by O'Regan et al. (1984). The importance of informativeness was observed in second fixations. They displayed an interaction such that the second fixation was always long if the first fixation had been forced to fall on an uninformative zone of the word. Also the summed initial fixation time showed a clear sensitivity to information distribution. However, the results of the synonymity judgment experiment are hard to interpret from the point of view of natural reading. The reason seems straightforward: One seldom faces the necessity of starting to read a word from its end.

On the other hand, the sentence comprehension experiments provided much clearer evidence of the influence of information upon the moment-to-moment control of eye guidance. In this experiment, the locus of fixation was not imposed upon the reader; instead, the reader was free to select the location of fixation or to choose not to fixate upon the word at all. It was found that inspection times were longer at the end of those words containing the information zone at the end. Again, this shows that fixation time can be influenced by the informativeness of a word part.

Experiments 2 and 3 confirmed that the beginning dominance phenomenon can be applied to continuous reading, too. That is, when reading long words, the reader tends to initially fixate the first part of the word regardless of its informativeness, the preferred landing position being the fourth or fifth letter of the word. On the other hand, the second half of the word is sometimes skipped, and it is skipped more often when it is redundant than when it is informative.

The data on continuous reading are also compatible with the data on reading words in isolation in that the first fixation duration was insensitive to information distribution within words, whereas the second fixation duration seemed sensitive to it.

References


**Appendix A**

**Materials Used in Experiment 1**

<table>
<thead>
<tr>
<th><strong>Information at the beginning</strong></th>
<th><strong>Information at the end</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Word used</strong></td>
<td><strong>Translation</strong></td>
</tr>
<tr>
<td>paärakennus</td>
<td>main building</td>
</tr>
<tr>
<td>nuorukainen</td>
<td>young man</td>
</tr>
<tr>
<td>pergamentti</td>
<td>parchment</td>
</tr>
<tr>
<td>vaarallinen</td>
<td>dangerous</td>
</tr>
<tr>
<td>meikäläinen</td>
<td>one of us</td>
</tr>
<tr>
<td>nykyäikainen</td>
<td>modern</td>
</tr>
<tr>
<td>pormestari</td>
<td>mayor</td>
</tr>
<tr>
<td>ylimääräinen</td>
<td>additional</td>
</tr>
<tr>
<td>elinympäräistö</td>
<td>surroundings</td>
</tr>
<tr>
<td>varsinainen</td>
<td>actual/real</td>
</tr>
<tr>
<td>jaanimurtaja</td>
<td>icebreaker</td>
</tr>
<tr>
<td>väkilannoitus</td>
<td>fertilization</td>
</tr>
<tr>
<td>monivuotinen</td>
<td>time over</td>
</tr>
<tr>
<td>kuraattori</td>
<td>counselor</td>
</tr>
<tr>
<td>lapsukainen</td>
<td>child</td>
</tr>
<tr>
<td>parlamentti</td>
<td>parliament</td>
</tr>
<tr>
<td>todennäköinen</td>
<td>probable</td>
</tr>
<tr>
<td>tavannukainen</td>
<td>customary</td>
</tr>
<tr>
<td>viranomaiinen</td>
<td>authority</td>
</tr>
<tr>
<td>eroottinen</td>
<td>erotic</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Word used</strong></th>
<th><strong>Translation</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>juoksuratta</td>
<td>track</td>
</tr>
<tr>
<td>leikkikalu</td>
<td>toy</td>
</tr>
<tr>
<td>sokeripala</td>
<td>sugar lump</td>
</tr>
<tr>
<td>maailmansota</td>
<td>world war</td>
</tr>
<tr>
<td>pakkasilmä</td>
<td>frost</td>
</tr>
<tr>
<td>perheenäiti</td>
<td>housewife</td>
</tr>
<tr>
<td>moottoritie</td>
<td>highway</td>
</tr>
<tr>
<td>keskennemeno</td>
<td>miscarriage</td>
</tr>
<tr>
<td>liikennevalo</td>
<td>traffic light</td>
</tr>
<tr>
<td>toimenpide</td>
<td>action</td>
</tr>
<tr>
<td>suurenmuslasilta</td>
<td>magnifying glass</td>
</tr>
<tr>
<td>kauppatori</td>
<td>market place</td>
</tr>
<tr>
<td>merkkivalo</td>
<td>light signal</td>
</tr>
<tr>
<td>henkilöauto</td>
<td>passenger car</td>
</tr>
<tr>
<td>pakkaslumi</td>
<td>frosted snow</td>
</tr>
<tr>
<td>moottorisaha</td>
<td>motor saw</td>
</tr>
<tr>
<td>linnunrata</td>
<td>milky</td>
</tr>
</tbody>
</table>

**Appendix B**

**Materials Used in Experiment 2**

The target word in each sentence is underlined, and the English translation is presented in parentheses.

**Story 1**

Lapsena Pekka asui kirkonkylässä. Han piti elinympäristöstään viehättävän. Kun maailmansota alkoi, kaikki tuhoutui. (In his childhood Pekka lived in a village. He found the surroundings very fascinating. When the world war broke out, everything was destroyed.)

**Story 2**

Maatalon isännällä oli paljon surua. Ensia moottorisaha löi hänen jalkansa. Sen jälkeen vaimo sai vielä keskenmenon. Pellon vakilannoitus jää myös tekemättä. (The farmer had much suffering. First the motor saw hit his leg. After that his wife had a miscarriage. So he didn't get his corn fields fertilized.)

**Story 3**

Juhlakokous oli juuri alkamassa. Kaikki parlamentin jäsenet istuutuivat. Puheenjohtaja otti esille käärön, joka oli erittäin vanhaa pergamentia. (The ceremonial meeting was about to begin. All the parliament members sat down. The chairman took out a package, which was of very old parchment.)

**Story 4**

Niimenen kaahasi pitkin moottoritietä. Pää hän huomasi poliisin merkkivalon. Hänä odotti varsinainen kuulesselu. Hänet tutkittiin kuin suurenmuslasilla. (Mr. Niimenen drove recklessly along the highway. Soon he noticed a light signal that policemen were showing. It was a real trial for him to wait. He felt that he was examined as thoroughly as if they had used a magnifying glass.)

**Story 5**

Asukkaiden monivuotinen haave toteutui. Kaupungintalon päärakennus valmistui. Asiaa juhlisti vielä pormestari puhe. Han totesi, että kyseessä oli merkkiteos, joka parantaa kauppatarin ilmettä. (The dream that the residents had had for many years came true. The main building of the city hall was completed. The mayor gave a ceremonial speech on the matter. He declared that the building stood as a monument, which made the marketplace surroundings impressive.)
**Story 6**


(The weather forecast says that the cold weather will continue. The sea-faring authorities are worried. All of the icebreakers are at work already. Therefore, we will have a normal winter.)

**Story 7**


(Drugs are a modern problem. Many measures have been proposed. People are especially worried about young men. This is emphasized by school counselors.)

**Story 8**

Edelläni ajanut henkilöauto pysähti, kun liikennevalo muutti punaiseksi. Tein vaaralliskiista hatalajarrutuksen. Ajorata oli nätä pakkasilmän peitossa.

(A car driving in front of me stopped when the traffic light turned to red. I made a dangerous emergency stop. The road was covered by frosty snow.)

**Story 9**


(The children of the house were playing in the yard. They were arguing about a toy. All the housewives were wondering when I would put the argument to an end.)

**Story 10**

Urheilukentän juoksurataa korjataan. Lisäksi rakennetaan ylimääräinen rata. Työ on todennäköisesti valmis tiistaina.

(The stadium track is being repaired. An additional track will be built. The work will probably be done by Tuesday.)

**Story 11**


(The moonshine reflected through the window pane to my eyes. I was just drinking tea with a sugar lump in my mouth. My thoughts were carried toward the Milky Way. The atmosphere acquired a slightly erotic touch.)

Received August 19, 1986
Revision received April 28, 1988
Accepted June 8, 1988