Word Ambiguity and the Optimal Viewing Position in Reading[⋆]

James J. Clark

Department of Electrical Engineering and Centre for Intelligent Machines,
McGill University
Montréal, Québec
Canada

J. Kevin O'Regan

Laboratoire de Psychologie Expérimentale
Centre National de la Recherche Scientifique
Université René Descartes
EPHE, EHESS, Paris, France

Abstract

The optimal viewing position phenomenon discovered by O'Regan *et al.* (1984) is characterized by a minimization of gaze duration on a word and maximization of word recognition rates when the eye fixates a word near its center. Subsequent studies (Holmes and O'Regan (1987), O'Regan and Lévy-Schoen (1987)) have shown that lexical structure can affect the location of the optimal viewing position. In this

Preprint submitted to Elsevier Preprint *9 June 1998*

paper we show that the optimal viewing position is near to the position which minimizes word ambiguity arising from incomplete recognition of the letters in the word. This conclusion is supported by a statistical analysis based on inter-letter correlations in English and French word corpuses.

Key words: eccentricity, fixation, reading, recognition, saccade

1 Introduction

O'Regan and coworkers (O'Regan *et al.* 1984, O'Regan and Lévy-Schoen 1987) discovered that there is an optimal position for the initial fixation on a word while reading. They found that the probability of refixation on the word as well as the gaze duration (the total time spent with the eye on a word) were minimized when the first fixation position of the eye on a word was near the center of the word. In addition, they observed that the probability of recognizing a briefly displayed word was greatest when the eye was fixated at this optimal position and decreased on either side of this point. Although the initial studies of O'Regan *et al.* involved situations where the word to be recognized was the first in a sequence of words, subsequent studies (McConkie *et al.* (1989), Vitu *et al.* (1990)) concentrating on recognizing words embedded in texts show that there is an optimal viewpoint in more general reading

---

⋆ Supported by the NISSAN Cambridge Basic Research Center

contexts. The effect is greatly reduced, however, for words in text as compared with words in isolation. In all studies, however, the location of the optimal viewpoint, though always near the center, varied somewhat depending on the situation (isolated words, reading) and on the response measure used (lexical decision, recognition probablity, fixation duration, refixation probability, etc.).

The most obvious possibility to explain the optimal viewing position effect is related to acuity limitations: because the acuity of the eye drops off dramatically on either side of the fixation point, even within the fovea (Levi, Klein and Aitsebaomo, 1985; Olzak and Thomas, 1986; Nazir, Heller and Sussman, 1992), it might be best to place the eye in such a way as to maximize the number of letters that are viewed with high acuity. Clearly, from considerations of symmetry, and as pointed out by O'Regan (1989), McConkie *et al.* (1988) and Nazir, O'Regan and Jacobs (1991), this hypothesis would predict that the optimal viewing position would be at the middle of the word. However, examination of the available literature suggests that the optimal viewing position may be asymmetrically placed, slightly left of the middle of words (O'Regan *et al.* 1984; Holmes and O'Regan, 1987; Brysbaert and d'Ydewalle, 1988; Vitu, O'Regan and Mittau, 1990; O'Regan and Jacobs, 1992; Vitu, 1993; Radach, 1993; Nazir, 1993; Farid and Grainger, 1996; Nazir, Jacobs and O'Regan, in press). Several mechanisms could contribute to this asymmetry: left/right differences in visual acuity (Nazir, 1991), hemispheric differences, perceptual learning (Nazir, 1993; Nazir, Jacobs and O'Regan, in press), lexical access strategies, and orthographic constraints.

It is difficult to tease apart the contributions of these different mechanisms without an explicit model of the processes underlying the optimal viewing position effect. As a step towards understanding these, the present paper investigates the contribution of one of the mechanisms, namely orthographic constraints. Orthographic constraints refer to the information contained in the letters of the word that have been identified. We make a simple assumption about what letters are recognized when the observer fixates at different positions in a word. Then we calculate an "ambiguity measure" which estimates how many words in the dictionary are compatible with the information available from a fixation at each possible fixation location in the word.

An interesting result transpires from our analysis: orthographic constraints are much stronger than expected, so much so that not only do they explain the asymmetry in the optimal viewing position curves, but they could actually by themselves explain the main component of the optimal viewing position effect, that is, the U-shaped dependence of recognition on fixation point. In other words, though previously the main possible mechanism underlying the optimal viewing position effect had been considered to be acuity drop off within the fovea and parafovea, the present results suggest that lexical constraints may actually suffice, or at least play an important role. The point is that under a pure "acuity" view, what determines the optimal viewing position is mainly the *number* of letters seen with high acuity. What our results show is that, in fact, it is not so much the *number* of letters, but rather the orthographic constraints imposed by *which* particular few letters of a word are directly fixated, and this may be mainly what determines the optimal viewing position effect. Further examinations of the effects predicted purely on the basis of orthographic constraints, in particular in regard to word length and word frequency, suggest that additional lexical access mechanisms at a more cognitive level must probably also be invoked.

It should be emphasized that our model is concerned with the ambiguity of the visual information regarding word identity that is presented to the recognition process at a given fixation. It says nothing about whether oculomotor processes actually move the eye to fixate at locations which minimize the ambiguity. A model which does this, using a entropy minimization analysis similar to the one given here, has been proposed by Legge *et al* (Legge 1997). Their system, called "Mr. Chips" does tend to generate fixation locations that minimize the uncertainty (or ambiguity) regarding the current word.

2 Estimating visual information: an approximation

Townsend, Taylor and Brown, (1971) showed that when the eye fixates a string of letters, the letter at the fixation location can be seen with near 100% accuracy, but the letters on either side of this location will be seen less well and will be confused with other letters. The set of confusible letters will depend on the exact type font and on the letter spacing and size, the other nearby letters, and the eccentricity of the letter (cf. Bouma (1970) who has estimated such confusion matrices). As we move further away from the fixated letter, the number of confusible letters increases. The end letters of words, however, will actually be seen better than letters that are inside the word. This is due to the phenomenon of "lateral masking", well known in the literature on letter and word perception.

In this paper we use an extreme simplification of the visual information gathering process, which we assume will give us an upper bound on the ambiguity; instead of attempting to tabulate all the letters that are confused with a given letter, we simply assume that the two letters that are nearest the fixation point can be seen correctly, but that all the other letters cannot be seen. An exception is made for the end letters of words which, because they are less subject to lateral masking, are assumed again to be seen perfectly. We expect that the use of this simplification will provide an upper bound to the ambiguity measure at each fixation location, since presumably in real reading more can be seen than just two letters and the end letters of words. We shall show however that even with this small amount of visual information, orthographic constraints are exceedingly strong! Furthermore, we shall see that the ambiguity depends on fixation location in a way which strongly resembles what is found for the optimal viewing position effect, thereby suggesting ambiguity as a possible mechanism underlying this effect.

3 The ambiguity measure

We wish to know, on average, how many words in the lexicon are compatible with the information that is obtained, using our simplified end-plus-two-interior-letter sampling scheme described above, when the eye fixates at different locations in a word. Suppose we consider the word "scattered", and assume that the eye is fixating beween the 3rd and 4th letters (i.e. between the a and the first t). Our sampling scheme thus assumes that what we observe is the pattern "s*at****d", where the *'s indicate letters that are not recognized. We can then ask, how many words in the lexicon are compatible with this pattern? A search shows that only the following three length 9 words in the Kucera and Francis corpus (Kucera and Francis, 1967) match this particular pattern: "scattered (27), spattered (2) and stationed (5)" (the quantities in parentheses are the relative frequency of occurence of the words in occurences per million words). Note that this corpus does not include such words as "shattered" and "statehood" which are also pattern matches and valid English words. We can repeat the search for pattern matches for all patterns of the form "$x_1 * x_2 x_3 * * * * x_4$". We can then compute the average, over all possible such patterns, of the number of pattern matches in a given corpus. This quantity is what we will refer to as the (unweighted) ambiguity measure (in this case for a fixation position of 3.5 with words of length 9). Appendix I gives details on how this ambiguity measure is calculated. An additional example of the dependence of word ambiguity with viewing position is given in figure 1.

FIGURE 1 GOES HERE

4 Statistical analyses of pattern ambiguity in English and French language wordlists

We studied our ambiguity measure for three different English language word corpuses and a French word corpus. These were the Webster's dictionary (210,680 words), the *ispell* dictionary (34,828 words), the Kucera and Francis (1967) word list (20,297 words), and the Trésor de la Langue Française (1971) (59,883 words). Words shorter than 5 letters, hyphenated words and words containing apostrophes were removed from these lists before the analysis was carried out. The Webster's word list is extensive and contains a large number of low frequency words, that is, words which are seldom encountered by readers. The *ispell* wordlist is that used by the *ispell* spelling checking program found on many Unix based computers. This wordlist contains a much higher proportion of high frequency words than the Webster's wordlist. The Kucera and Francis and the Trésor de la Langue Française wordlists have the advantage of also including relative frequencies (occurrences per million words) for each of the words. For analyses carried out with these wordlists we are able to use the relative word frequency information to compute a weighted ambiguity measure.

FIGURE 2 GOES HERE

FIGURE 3 GOES HERE

FIGURE 4 GOES HERE

FIGURE 5 GOES HERE

The results of the statistical analyses on the four wordlists are displayed in figures 2 through 5. The curves all have the same characteristic convex bowl shape, irrespective of the word length, the wordlist used, and the frequency of the words in the wordlists. For each curve there is a clearly defined minimum. Note the surprising fact that, even for the very large Webster's wordlist, the average ambiguity near the word centers was less than 2.0 interpretations per pattern. This was true even for words of 11 letters. This low level of ambiguity indicates that knowledge of an interior letter pair, if optimally located, along with knowledge of the word length and the first and last letters of a word provide powerful constraints on the identity of the word. Contextual constraints from neighboring words presumably reduce this ambiguity level even further.

FIGURE 6 GOES HERE

The location of the fixation that results in the minimum ambiguity as a function of word length is plotted in figure 6. This location was determined via interpolation by fitting a parabola to the curves and computing the minimum of the parabola. The location of the minimum closely tracks a position half a letter to the left of the center of the word for all the wordlists used. This is clearly seen in figures 2 through 5 as well. This is an important feature, and is also observed in the human data. We will discuss this later in the paper.

An interesting aspect of the ambiguity curves concerns the effect of word length. If we plot the value of the minimum ambiguity near the middle of the word as a function of word length, as done in Figure 7, we see that ambiguity increases with word length up to a maximum value for word lengths of 8 to 10 letters, depending on the corpus, and then decreases again. This behavior is a consequence of the combined action of two opposing factors, the weaker constraint imposed by the end letters on letters inside long words, making them less ambiguous, and the fact that there are more long words than short words. The point where the ambiguity starts dropping depends on the size of the corpus, and is later for larger corpuses. The curves shown in figure 7 are for the case for fixation at the optimal location. As the fixation location deviates from the middle, the cross-over point moves towards shorter word lengths. This can be seen in figures 2-5, where the ambiguity curves cross over each other so that for fixation locations near the ends of the words the ambiguity is less for long words than for short words. One could argue that this result may be an artifact caused by, for short words, the letters visible from the fixation location overlapping the end letters, effectively reducing the number of visible letters. In our analyses, however, we intentionally restricted fixation locationss to those that did not overlap the end letters.

FIGURE 7 GOES HERE

Figures 8 and 9 show the ambiguity of words in the Kucera and Francis and in the Trésor de la Langue Française wordlists separately for high and low frequency words. For compatibility with empirical optimal viewing position studies we have chosen word frequency classes of 50- 500 per million for the high and 0.5-5 per million for the low frequency categories. The ambiguity for words in the high and the low frequency classes was calculated by obtaining the average of the number of words *in the whole corpus* that are visually compatible with each word in the high or low frequency class, respectively. The low frequency words have flatter ambiguity curves, and these lie below the curves for the high frequency words: letter patterns in low frequency words are less ambiguous than those in high frequency words. Our intuition that high frequency words share a smaller set of common letter groupings than low frequency words seems to be confirmed by this finding. Interestingly however, empirical optimal viewing position curves show exactly the opposite trend, as seen by comparing figures 8 and 9 with figure 10.

FIGURE 8 GOES HERE

FIGURE 9 GOES HERE

4.1 Comparison with empirical optimal viewing position curves

The empirical data on the optimal viewing position phenomenon shows a number of characteristics which can compared with the results of the statistical calculations performed on the word corpuses.

Before examining the relationship of our statistical data to the empirical optimal viewing position data, we must choose which empirical data to use as a basis of comparison: indeed, the empirical data available in the literature differ quite strongly in the degree of asymmetry, the word length and the word frequency effects. It may be that the results obtained depend on factors such as the language used, the task used to measure the effect (lexical decision, naming latency, normal reading) and the measure being used to determine the effect (number of eye fixations, probability of refixation, total gaze duration, probability of correct recognition, etc.). Nazir (1991, 1993) has argued that, when eye movements are used to measure the optimal viewing position, low-level oculomotor mechanisms come into play that may affect the measured variables in a way which is not directly dependent on word recognition processes: the eye may refixate or increase its fixation durations for reasons related to oculomotor scanning and not to visual word processing. Furthermore, as noted by Vitu *et al.* (1990), tasks such as continuous text reading may involve global eye movement strategies, parafoveal preprocessing, as well as text comprehension components that mask or modify the recognition processes involved in recognizing each word. For these reasons, in addition to optimal viewing position data obtained through eye movement measurements in normal reading, it may be reasonable to look at data obtained on isolated words using the lexical decision task in which eye movement characteristics are not used as dependent variables. But even these data should be considered with caution, because the particular nonwords used in the word/non-word decision task may have induced particular response strategies.[1] Finally, it is important to note that in all the studies on the optimal viewing position, only a small set of words will ever have been tested. The optimal positions for these subsets of the lexicon may differ considerably from the means that would be obtained if all the words in a lexicon had been tested. In fact, Holmes and O'Regan (1992) and Farid and Grainger (1996) expressly manipulated the morphological structure of words and their orthographic constraints and found systematic deviations in the location of the optimal viewing position. Overall, therefore,

---

[1] The time to accomplish a lexical decision task is, of course, less directly linked to ambiguity than would be a word recognition task, since, depending on the particular word and non-word sets used, response strategies might be used by subjects which do not require full word recognition. However, in the particular experiments considered here, the incidence of such strategies was probably reduced by the fact that nonwords only differed from existing words by a change in a single letter or in two letters. We also feel justified in using this measure because in the literature on word recognition, lexical decision is generally highly correlated with other measures such as naming latency and recognition accuracy.

it will be best to compare our ambiguity data with a variety of empirical studies of optimal viewing position data: those obtained using lexical decision on isolated words without eye movement measures (O'Regan and Jacobs (1992, experiment 1); Nazir (1993); Nazir *et al.* (in press)), as well as those using eye movement on isolated word reading (O'Regan *et al* (1984); Vitu *et al* (1990)), and on text reading (McConkie *et al.* (1989); Vitu *et al.* (1990); Vitu (1991)).

4.1.1 U-Shape of optimal viewing position curves

FIGURE 10 GOES HERE

The most striking characteristic of optimal viewing position data, whether they are obtained from reading isolated words or normal text, and whether lexical decision or eye movement parameters are measured, is that they are U-shaped. A typical example is plotted in Figure 10, taken from experiment 1 of O'Regan and Jacobs (1992), which measured lexical decision time on isolated words. McConkie et al. (1988), O'Regan (1989), Nazir (1991), and Nazir, O'Regan and Jacobs (1991) all attributed this U-shapedness to the acuity drop-off within the parafovea, that is, the fact that at the middle of words, more letters are seen with high acuity. It is particularly interesting that the curves of ambiguity we have obtained from the statistical analysis of the word corpuses are also U-shaped, but not because more letters are seen, rather because more informative letters are seen. This point is relevant to a study performed by Nazir, Jacobs and O'Regan (in press) using "butterfly" shaped configurations in which the size of the letters of a word around the current fixation point were magnified so that the letters further away from the fixation point were actually more visible than the ones that were being directly fixated. The authors observed that this manipulation removed the optimal viewing position effect for words of length five letters, as expected from the view that the number of letters seen with high acuity is what limits word recognition. Interestingly however, for words of length 9 letters the butterfly manipulation had no effect; thus, as for normal shaped nine letter words, butterfly-shaped are harder to recognize from a fixation position near the beginning or near the end of the word, even when the other letters are increased in size so that they were guaranteed to be perfectly visible. This result is difficult to understand from a pure acuity based model of the optimal viewing position effect, and the authors were led to invoke a "perceptual learning" explanation in which they proposed that words are best recognized from a viewing location at which they are habitually fixated – the idea is that since statistically the eye tends to land nearer the beginning of words, through perceptual learning this comes to be the place where it is best to fixate. On the other hand, the present results give a natural alternative explanation to the data of Nazir *et al.* If we assume that words of length five letters are a special case, since all but one of the letters will be perfectly visible, then the result is exactly what would be predicted from a model such as the one employed here, in which the information obtained from the word was effectively independent of acuity, and depended only on the orthographic constraints provided by the particular letters that were directly fixated (and the end letters of the word).

While the present four-letter sampling model of recognition accounts for what would otherwise be a puzzling result in the case of the butterfly-shaped nine-letter words, other results show that acuity probably must be playing a role. Nazir, Heller and Sussman (1992) manipulated the spacing between the letters of words and observed that the shapes of the optimal viewing position curves for seven and nine-letter words were modified in a way which was predictable from acuity considerations.

In summary, whereas previously the U-shape of the optimal viewing position curves was assumed to be caused by the way acuity drop-off affects the *number* of letters that can be seen, the present analyses show that it may be not the number but the *choice* of letters that are fixated directly that plays the essential role in making the middle region of the word the most advantageous for fixation. Extensions of the present simple model, in which more than just two letters are assumed to be visible, or in which letter confusion probabilities are taken into account, or finally in which word length information is assumed to be less reliably known, may provide a significant improvement over the pure acuity-based explanation of the optimal viewing position effect.

4.1.2 Asymmetry of optimal viewing position curves

FIGURE 11 GOES HERE

As seen in Figure 10, the optimal viewing position curves are asymmetric, with their minima systematically displaced left of the center of words. In addition, the slopes of the left and right branches of the curves are different, with the right branch being steeper. Both these aspects of the curves are precisely what is found in our curves of the statistical ambiguity measure.

For better comparison of the empirical curves to statistically measured ambiguity measures, Figure 11 plots the interpolated position of the minimum of the optimal viewing position for the O'Regan and Jacobs (1992, Expt 1) as well as the four other experiments (Nazir *et al.* (1991), lexical decision task with isolated word; McConkie *et al.* (1988), refixation probability in text reading; Vitu *et al.* (1990), refixation probability in isolated words and in text reading). Depending on the study, the exact position of the optimum varies somewhat, but in all cases, it tends to be a fairly constant amount to the left of the middle of the words, just as in the statistical ambiguity measures. Only in McConkie et al.'s (1988) study was the optimal position slightly to the right of center.

It would appear, then, that the left/right asymmetry for optimal viewing position curves that would be predicted purely by invoking orthographic constraints as a mechanism would have many similarities to the empirically observed data, which is generally to the left of center, but to different degrees.

As noted above, several factors may contribute to the variability of the empirical data: the reading situation (isolated words or continuous reading), the response measure used – lexical decision, naming latency, eye movements), the subset of words used for the tests. Further work is necessary to see whether such factors, when added to the simple four-letter coding hypothesis considered here, suffice to explain the data patterns. However it may be that other factors will also have to be incorporated into a more sophisticated model: left/right asymmetry in acuity, hemispheric differences in the brain, left-to-right processing strategies in lexical access, processes of morphological decomposition, and perceptual learning.

4.1.3 Effects of word frequency

For word frequency a very clear difference is observed between the empirical optimal viewing position curves and the ambiguity curves: whereas the empirical optimal viewing position data show that low frequency words are harder to recognize than high frequency words, the ambiguity data show the opposite.

The differences between the empirical word recognition data and our ambiguity analysis as concerns the effects of word frequency show that factors other than orthographic constraints are at work in determining empirical lexical decision latency and time to recognize a word. Most models of word recognition (cf e.g. Carr and Pollatsek, 1985; Jacobs and Grainger, 1994) assume that letters and words having high frequencies of occurrence receive more weight in the decision process (or correspond to processing units having lower firing thresholds) than letters and words that occur infrequently. This idea is intuitively plausible and reasonable in the context of a learning device, where there will be more exposure to frequent letters and words. Furthermore the idea is coherent with the empirical finding that frequent and short words are easier to recognize. But the present statistical calculations show that in fact a Bayesian decision device might not actually have these properties. It would be interesting to evaluate current word recognition models with this in mind.

4.2 Double Fixations

The longer a word, the greater the probability that it will receive more than one fixation. O'Regan (1992) has postulated that the second fixation becomes necessary *because* of the initial fixation being away from the optimal viewing position. It is conceivable, however, that the eye is capable of judging the length of a word before it fixates it. The eye could then use this information to decide whether a single fixation is sufficient or whether two fixations will be needed to reliably recognize the word. If it decides the latter, then it will saccade to a location inside the word which is optimal for the two-fixation case.

What would be the optimal locations for these fixations? A reasonable strategy is to fixate the word in two different locations spaced apart and in the interior of the word. While data on the optimality of the locations of multiple fixations (in terms of maximizing recognition performance or minimizing gaze durations) does not appear to have been published, some studies have noted the actual locations of the first and second fixations on words that are multiply fixated (e.g. O'Regan (1992)). Here it is generally found that when two fixations occur on a word, they are distributed in a spaced-out way in the word, that is, if the first is at the beginning, then the second is at the end, and vice versa. Is this behavior what would be expected from lexical statistics?

To determine what the optimal viewing positions are for the two-fixation case, using our criterion of minimal lexical ambiguity, we performed another parametric statistical study. In this study we tabulated the average number of matches to letter patterns of the form "$x_1 * * x_2 x_3 * * x_4 x_5 * * x_6$", where the locations of the two interior letter pairs were variable parameters. The results are summarized in figures 12 through 14. In order to collapse the 2-dimensional parameter space into a form suitable for displaying in 1-dimensional curves, we plot the minimum ambiguity over either the first fixation location or the second fixation location.

FIGURE 12 GOES HERE

FIGURE 13 GOES HERE

FIGURE 14 GOES HERE

The results are qualitatively similar to the one-fixation case, in that the average ambiguity versus fixation position curves are convex with a clearly defined minimum. The location of the minima as a function of word length can be seen in the graph shown in figure 14. Unlike the one-fixation case the optimal position of the first fixation is *not* near the center of the word, but is shifted strongly to the left. Likewise the optimal position of the second fixation is shifted to the right of the center of the word. Note that the slope of the second fixation location vs. word length curve is the same as that of the word center vs. word length line, indicating a true shift of about one letter to the right of the center of the word. In contrast to this, the slope of the first fixation location vs. word length curve is somewhat flatter with the optimal first fixation location lying near the 4th letter for all word lengths tested.

There is an interesting piece of experimental evidence which may lend support to this analysis. This comes from a study performed by Vitu (1991), who examined the location of the first fixation made by the eye on a word with the aim of determining whether there is a center-of-gravity effect in reading. The center-of-gravity effect (see for example, Coren and Hoenig (1972) or Findlay (1982)) refers to the tendency of the eye to move to the center-of-gravity of a stimulus array when saccadic latencies are very short. Vitu's idea is that the reason the eye moves to the center of a word is not that it is optimal for some reason, but that is attracted there due to the center-of-gravity effect. In the first experiment described in Vitu (1991) there is a shift (roughly following the center of the word) in the landing position going from 5 letter words to 9 letter words, but there is no shift from 9 to 13 letter words, staying on the average at a location between the 3rd and 4th letters. This result can be explained by our theory as it predicts a linear increase in the shift of landing position with word length up to 9 letters or so. For longer words, one could hypothesize that more than one fixation will be made. In this case the optimal landing position is seen to lie around 3.5 to 4 letters and is only weakly dependent of word length. This suggests that the observed fixation patterns observed in Vitu's study might be due not to the center-of-gravity effect, but to the eye seeking out the optimal initial viewing position for multiple fixations.

5 Discussion

In this paper we have proposed an explanation of the optimal viewing position effect observed by O'Regan et al. Instead of appealing to the idea that the effect is related to the fact that fixating near the middle of words allows the maximum number of letters to be seen with the maximum visual acuity, our explanation is instead based on orthographic constraints. Assuming that only a few letters of a word are used to establish the identity of a word, because of orthographic constraints, surprisingly reliable recognition can usually be obtained if the eye is placed just left of the middle of the word.

But then what happens when the word is not uniquely identifiable? The different word candidates compatible with the sampled information may compete with each other in the way proposed in many connectionist models of word recognition (e.g. Grainger and Jacobs, 1996). In this process, contextual constraints deriving from the semantic and syntactic context of the word, as well as possible parafoveal preview information obtained about the word, may help one word more than the others, and recognition may ultimately ensue, although after a delay depending on the ambiguity of the information available.

Another possibility is that an eye movement may occur, leading the eye to another position in the word, from which further information can be sampled, allowing the word to be disambiguated. Legge et al. (1998) have modeled the refixation process from an ideal observer point of view similar to the view taken here, and shown that eye movement patterns generated by such a scheme resemble those that are found in real reading. However, the sampling scheme Legge et al. used did not involve the particular end-letter and length information code we have assumed in our model. Nonetheless, such ideas are compatible with a number of findings (eg. McConkie *et al.* (1988); Vitu *et al.* (1990)) showing that there is a greater tendency for the eye to make a second fixation in a word, as its initial location in the word deviates from the optimal, (just left of) center position. The influence of contextual information is also suggested by Vitu *et al.*'s finding that the probability of refixation on a word was much less (25% vs. 75%) when the word was embedded in text than when the word was isolated. The model proposed here is of course an extreme simplification. Subjectively, when you fixate in a long word, you feel you can clearly see two or three letters at the fixation point, the end letters, and additionally some other cues in the word, such as the presence of distinctive letters like "o" or "l", and the presence of ascending and descending elements. Thus our estimation here of the information available from a simple four-letter code is certainly a significant underestimation of the available information: undoubtedly in real word recognition, people use more than the two letters that they are directly fixating, the length of the word, and the end letters of the word. What is extremely surprising, then, is that even assuming only this very small amount of information, the word can very often be uniquely identified, provided fixation is placed at the optimal location.

Word recognition is no doubt a much more complex operation than our model implies, and depends on many factors which we do not address. Many of these factors could affect word recognition performance in a fixation location dependent manner, and would thus influence the optimality of a viewing position. The simple 4-letter coding model we have proposed is not meant to be a realistic word recognition scheme, but an upper bound to such a scheme. The fact that it predicts many aspects of the optimal viewing position curves shows that the othographic constraint principle underlying our scheme may be a strong contributor to the optimal viewing position effect. It is clear however that a more realistic word recognition model would incorporate the sampling of letter information from more than just the four letters considered here, and would additionally involve disambiguation mechanisms making use of contextual constraints. The model presented in this paper is another step along the road of quantifying the effect of within-word lexical constraints on word recognition.

References

[1] Bouma, H. (1970), "Interaction effects in parafoveal letter recognition", *Nature*, Vol. 226, pp 177-178

[2] Bouma, H. (1971), "Visual recognition of isolated lower-case letters", *Vision Research*, Vol. 11, pp 459-474

[3] Bouma, H. (1973), "Visual interference in the parafoveal recognition of initial and final letters of words", *Vision Research*, Vol. 11, pp 459-474

[4] Brysbaert, M. and d'Ydewalle, G. (1988), "Callosal transmission in reading", in **Eye Movement Research: Physiological and Psychological Aspects** G. Luer, U. Lass, and J. Shallo-Hoffman (eds.), Göttingen: Hogrefe, pp 246-266

[5] Carr, T.H. and Pollatsek, A. (1985) Recognizing printed words: A look at current models. In D. Besner, T.G. Walker, and G.E. MacKinnon (Eds.) **Reading research: Advances in theory and practice 5** San Diego, Academic Press, pp 1-82

[6] Coren, S. and Hoenig, P. (1972), "Effect of non-target stimuli upon length of voluntary saccades", *Perceptual Motor Skills*, Vol. 34, pp 499-508

[7] Farid, M. and Grainger, J. (1996), "How initial fixation position influences visual word recognition: A comparison of French and Arabic", *Brain and Language*, Vol. 53, No. 3, pp 351-368

[8] Findlay, J.M. (1982), "Global visual processing for saccadic eye movements", *Vision Research*, Vol. 22, pp 1033-1045

[9] Grainger, J. and Jacobs, A.M. (1996), "Orthographic processing in visual word recognition: A multiple read-out model", *Psych. Review*, Vol. 103, No. 3, pp 518-565

[10] Holmes, V.M., and O'Regan, J.K. (1987), "Decomposing French words", in J.K. O'Regan and A. Lévy-Schoen (eds.) **Eye Movements: from physiology to cognition**, North Holland, Amsterdam, pp 459-466

[11] Holmes, V.M., and O'Regan, J.K. (1992), "Reading derivationally affixed french words", *Language and Cognitive Processes*, Vol. 7, pp 163-192.

[12] Kucera, H., and Francis, W.N. (1967), **Computational analysis of present-day American English**, Brown University Press, Providence, RI, USA.

[13] Legge, G.E., Klitz, T.S., and Tjan, B.S. (1997), "Mr. Chips: An ideal-observer model of reading", *Psych. Review*, Vol. 104, No. 3, pp 524-553

[14] Levi, D.M., Klein, S.A. and Aitsebaomo, A.P. (1985), "Vernier acuity, crowding and cortical magnification", *Vision Research*, Vol. 25, pp 963-977

[15] McConkie, G.W, Kerr, P.W., Reddix, M.D., and Zola (1988), "Eye movement control during reading: I: The location of initial eye fixations on words", *Vision Research*, Vol. 28, pp 1107-1118

[16] McConkie, G.W, Kerr, P.W., Reddix, M.D., Zola, D., and Jacobs, A.M. (1989), "Eye movement control during reading: II: Frequency of refixating a word", *Perception and Psychophysics*, Vol. 46, pp 245-253

[17] Nazir, T.A. (1991), "On the role of refixations in letter strings: the influence of oculomotor factors", *Perception and Psychophysics*, 49, pp 373-389

[18] Nazir, T.A. (1993), "On the relation between the optimal and the preferred viewing position in words during reading", in **Perception and Cognition: Advances in Eye Movement Research**, G. d'Ydewalle and J. van Rensbergen (eds.), Amsterdam: North-Holland, pp 349-361

[19] Nazir, T.A., Heller, D., and Sussman, C. (1992), "Letter visibility and word recognition: The optimal viewing position in printed words", *Perception and Psychophysics*, Vol. 52, pp 315-328

[20] Nazir, T.A., Jacobs, A.M., and O'Regan, J.K. (1998), "Letter legibility and visual word recognition", *Memory and Cognition*, in press.

[21] Nazir, T.A., O'Regan, J.K., and Jacobs, A.M. (1991), "On words and their letters", *Bulletin of the Psychonomic Society*, Vol. 29, pp 171-174

[22] Olzak, L.A. and Thomas, J.P. (1986), "Seeing spatial patterns", in **Handbook of Perception and Human Performance V. II**, K.R. Boff, L. Kaufman, and J.P. Thomas (eds), New York: Wiley, pp 7:1-7:56

[23] O'Regan, J.K. (1989), "Visual acuity, lexical structure, and eye movements in word recognition", in **Working Models of Human Perception**, B. Elsendoorn and H. Bouma, (eds), London: Academic Press, pp 261-292

[24] O'Regan, J.K. (1992), "Optimal viewing position in words and the strategy-tactics theory of eye movements in reading", in **Eye Movements and Visual Cognition: Scene Perception and Reading**, K. Rayner (ed.), New York: Springer-Verlag, pp 333-354

[25] O'Regan, J.K. and Jacobs, A.M. (1992), "Optimal viewing position effect in word recognition: a challenge to current theory", *Journal of Experimental Psychology: Human Perception and Performance*, Vol. 18, pp 185-197

[26] O'Regan, J.K. and Lévy-Schoen, A. (1987), "Eye movement strategy and tactics in word recognition and reading", in M. Coltheart (ed.), *Attention and Performance XII: The psychology of reading*, pp 363-383, Erlbaum, Hillsdale NJ.

[27] O'Regan, J.K., Lévy-Schoen, A., Pynte, J. and Brugaillère, B. (1984), "Convenient fixation location within isolated words of different length and structure", *Journal of Experimental Psychology: Human Perception and Performance*, 10(2), pp 250-257

[28] Radach and Kempe, (1993), "An individual analysis of initial fixation position in reading", in **Perception and Cognition: Advances in Eye Movement Research**, G. d'Ydewalle and J. van Rensbergen (eds.), Amsterdam: North-Holland, pp 213-225

[29] Seidenberg, M.S., Tannehaus, M.K., Leiman, J.M., and Bienkowski, M. (1982), "Automatic access of the meanings of ambiguous words in context: some limitations of knowledge-based processing", *Cognitive Psychology* Vol. 14, pp 439-537

[30] Townsend, J.T., Taylor, S.G., and Brown, D.R. (1971), “Lateral masking for letters with unlimited viewing time”, *Perception and Psychophysics*, Vol. 10, pp 375-378

[31] Vitu, F. (1991), “The existence of a center of gravity effect during reading”, *Vision Research*, Vol. 31, pp 1289-1313

[32] Vitu, F., O’Regan, J.K. and Mittau, M. (1990), “Optimal landing position in reading isolated words and continuous text”, *Perception and Psychophysics*, Vol. 47, pp 583-600

6 Appendix I: Statistical Measures of Ambiguity

Our hypothesis as to the source of the optimal viewing position effect states that the optimal viewing position is that which minimizes the ambiguity due to the partial letter identity information available at the fixation location. In this appendix we quantify the notion of ambiguity, and provide statistical measures of ambiguity that can be used in empirical analyses of lexical structure.

Our model of word recognition supposes that readers recognize the identity of letters at the ends of words and a number of letters at a fixation location in the interior of the word. The identities of the remaining letters in the word are assumed to be left undetermined. We further assume that the fixation location is such that the range of letters recognized at fixation do not overlap the end letters of the word. The word length is also assumed to be known. Thus, for words of length L the interior fixation location can lie in the range $x \in [1+(s+1)/2, L-(s+1)/2]$, where $s \leq L-2$ is the number of letters recognized at the fixation location. For a given word length L and number of interior letters recognized s, each fixation location x admits a set of letter patterns, $\mathcal{P}(x; L, s)$. For example if $L = 4$ and $s = 1$ there are two possible fixation locations and two sets of letter patterns: $\mathcal{P}(3; 4, 1) = \{$(aa*a),(aa*b),...,(zz*z)$\}$ and $\mathcal{P}(2; 4, 1) = \{$(a*aa),(a*ab),...,(z*zz)$\}$, where $*$ indicates that any single letter can occur in this location and match the pattern.

For a typical extensive list of words, or dictionary $\mathcal{D}$, there will be many patterns $p \in \mathcal{P}(x; L, s)$ which match multiple words w in the wordlist. Such patterns are therefore ambiguous. As a measure of ambiguity we can use the average number of matches per pattern.

The average number of matches per pattern could be computed by counting the number of words matching a given pattern and then averaging this value over all patterns found in the lexicon (patterns not found in the lexicon will obviously have no words that match it). In a straightforward implementation this would require that every word in the lexicon be matched against every pattern, which can be a computationally expensive exercise. Fortunately, a less computationally demanding method can be used, as follows. Let $N(p_i)$ be the number of words $w \in \mathcal{D}$ which match a pattern $p_i \in \mathcal{P}(x; L, s)$. Define N_p to be the number of patterns in the set of patterns $\mathcal{P}(x; L, s)$. Clearly, $N_p = 26^{2+s}$ for all values x and L, since the identity of $2 + s$ letters are specified in each pattern. Let N be the total number of words of length L in the dictionary $\mathcal{D}$. It is seen that $N = \sum_{i=1}^{N_p} N(p_i)$, since the patterns are mutually exclusive. The average number of matches is then given by $A = \frac{N}{N_p}$. This expression requires only counting the number of letters of length L in the lexicon, and enumerating the number of distinct patterns. This can be done in a single pass over the file, and was actually implemented using a Unix shell script, rather than a C program.

Figure Captions

Figure 1. Illustration of a simple approximation of what the eye can see when it fixates, A: near the middle of the word “undergraduate”, and B: nearer to the beginning of the word. On the right are shown the words in an english lexicon (Webster’s dictionary) that are compatible with the visual information assumed to be available. The example given here is a word having a complex morphological structure. The morphology of a word is correlated with the statistics of the lexicon, and will influence the number of candidates that are generated by our coding scheme. However in our model, we assume no independent morphological parsing mechanism.

Figure 2. The ambiguity for patterns defined by the first letter, last letter and an interior letter pair, as a function of the position of the first letter in the letter pair. The different curves represent different word lengths from 5 to 11 letters. The ispell word list was used to generate this graph.

Figure 3. The ambiguity for patterns defined by the first letter, last letter and an interior letter pair, as a function of the position of the first letter in the letter pair. The different curves represent different word lengths. The Webster’s word list was used to generate this graph.

Figure 4. The ambiguity for patterns defined by the first letter, last letter and an interior letter pair, as a function of the position of the first letter in the letter pair. The different curves represent different word lengths. The Kucera

and Francis word list was used to generate this graph.

Figure 5. The ambiguity for patterns defined by the first letter, last letter and an interior letter pair, as a function of the position of the first letter in the letter pair. The different curves represent different word lengths. The Trésor de la Langue Française word list was used to generate this graph.

Figure 6. The (interpolated) location of the first eye fixation that yields the minimum average ambiguity in the case of recognizing interior letter pairs. The fixation location is taken to be halfway between the letters in the letter pair. Shown are the results as a function of word length for the ispell, Webster's and Kucera and Francis (with and without frequency weighting) word lists.

Figure 7. The value of the minimum ambiguity as a function of word length.

Figure 8. The ambiguity curves for the Kucera and Francis wordlist separated into high and low frequency categories.

Figure 9. The ambiguity curves for the French wordlist separated into high and low frequency categories.

Figure 10. Data from O'Regan and Jacobs (1992) showing the time it takes to decide whether a string of characters appearing at different positions relative to the eyes' fixation point is a word or not. In the experiment, the eye fixated in the gap between two vertically aligned line segments in the middle of the screen. After 500 ms the lines disappeared and the test word or nonword

appeared. The word or nonword was laterally displaced in relation to the gap in such a way that when it appeared, the eye was fixating one of five positions defined over the length of the stimulus. Curves are shown for the recognition of words of low and high frequency of occurrrence in the language, and each group of curves corresponds to words of length 5, 7, 9, and 11 letters.

Figure 11. The interpolated position of the location where ambiguity is the smallest, or for the empirical studies, where the dependent variable (percent correct, lexical decision time, refixation probability, etc.) is at its minimum, as a function of word length.

Figure 12. The minimum (over locations of the second fixation) average number of interpretations of a pattern given two fixations of the eye on the word, as a function of the location of the first fixation. At each fixation a pair of letters is recognized. In addition, the first and last letters of the word are assumed to be known. The Webster's word list was used to generate this graph.

Figure 13. The minimum (over locations of the first fixation) average number of interpretations of a pattern given two fixations of the eye on the word, as a function of the location of the second fixation. At each fixation a pair of letters is recognized. In addition, the first and last letters of the word are assumed to be known. The Webster's word list was used to generate this graph.

Figure 14. The (interpolated) location of the first and second eye fixations that yield the minimum average ambiguity in the case of recognizing pairs of

interior letter pairs.

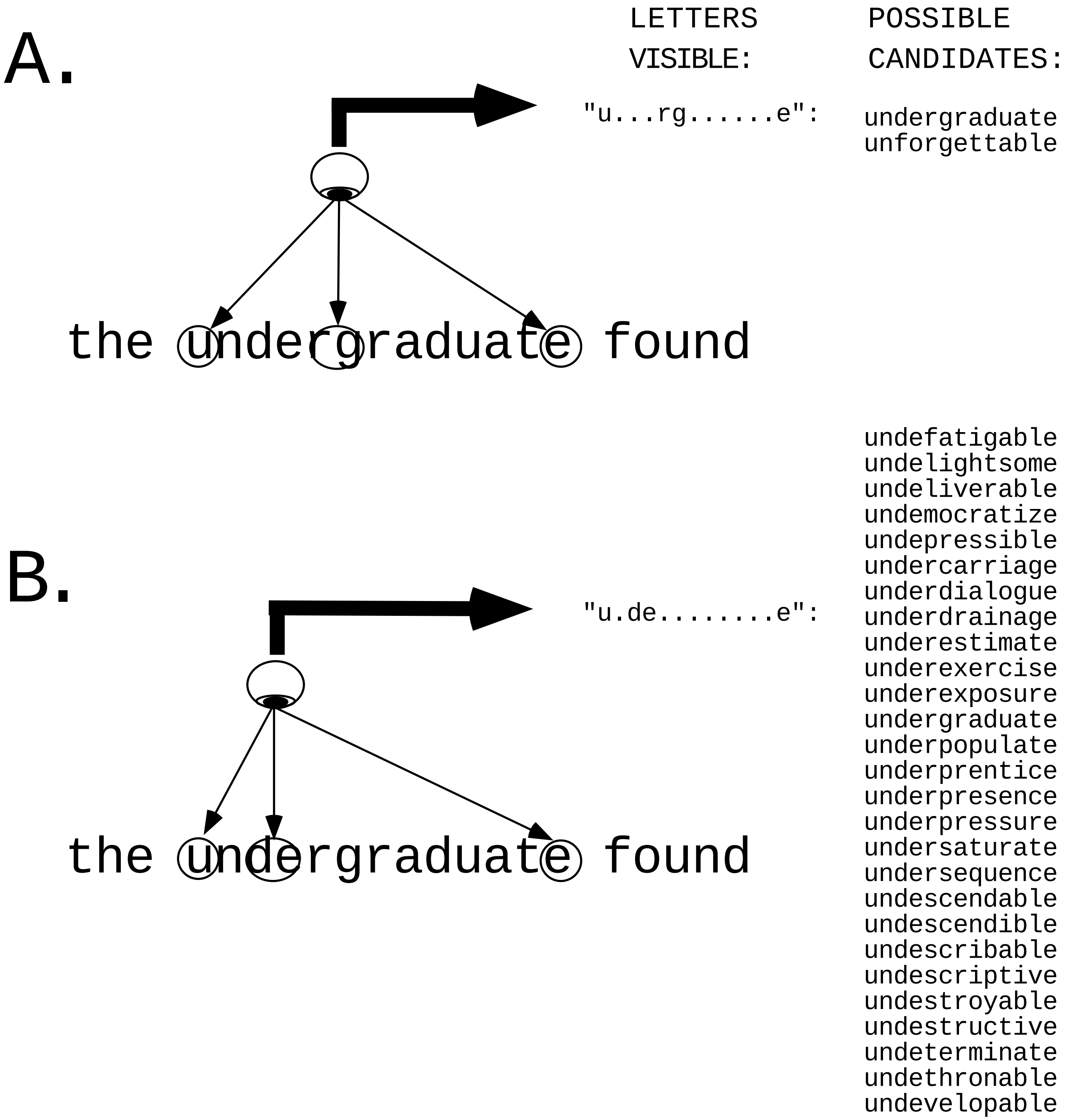

Figure 1. Illustration of a simple approximation of what the eye can see when it fixates, A: near the middle of the word "undergraduate", and B: nearer to the beginning of the word. On the right are shown the words in an english

lexicon (Webster's dictionary) that are compatible with the visual information assumed to be available. The example given here is a word having a complex morphological structure. The morphology of a word is correlated with the statistics of the lexicon, and will influence the number of candidates that are generated by our coding scheme. However in our model, we assume no independent morphological parsing mechanism.

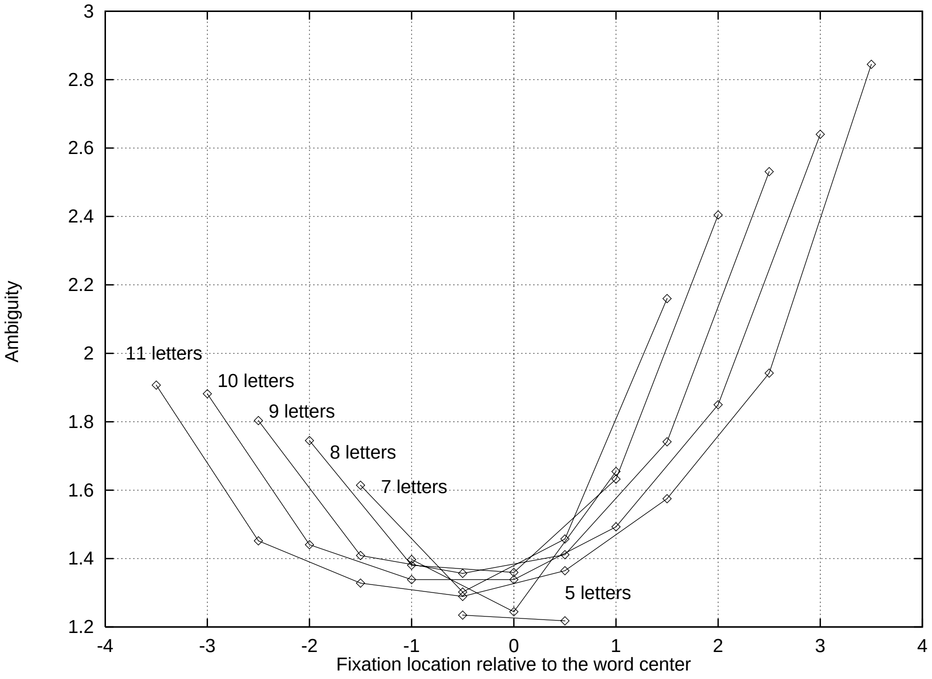

Figure 2. The ambiguity for patterns defined by the first letter, last letter and an interior letter pair, as a function of the position of the first letter in the letter pair. The different curves represent different word lengths from 5 to 11 letters. The ispell word list was used to generate this graph.

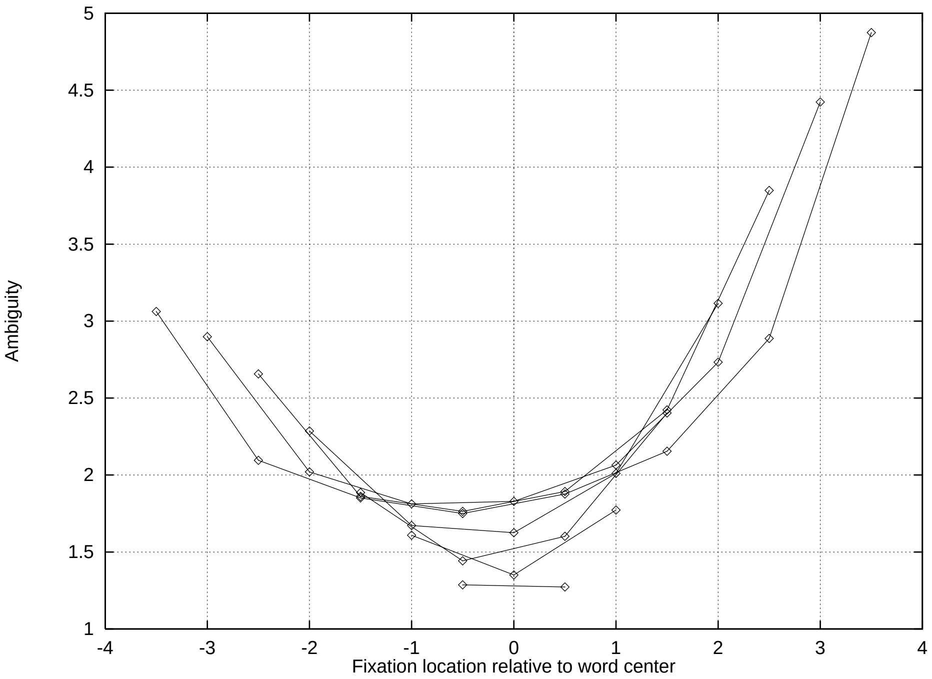

Figure 3. The ambiguity for patterns defined by the first letter, last letter and an interior letter pair, as a function of the position of the first letter in the letter pair. The different curves represent different word lengths. The Webster's word list was used to generate this graph.

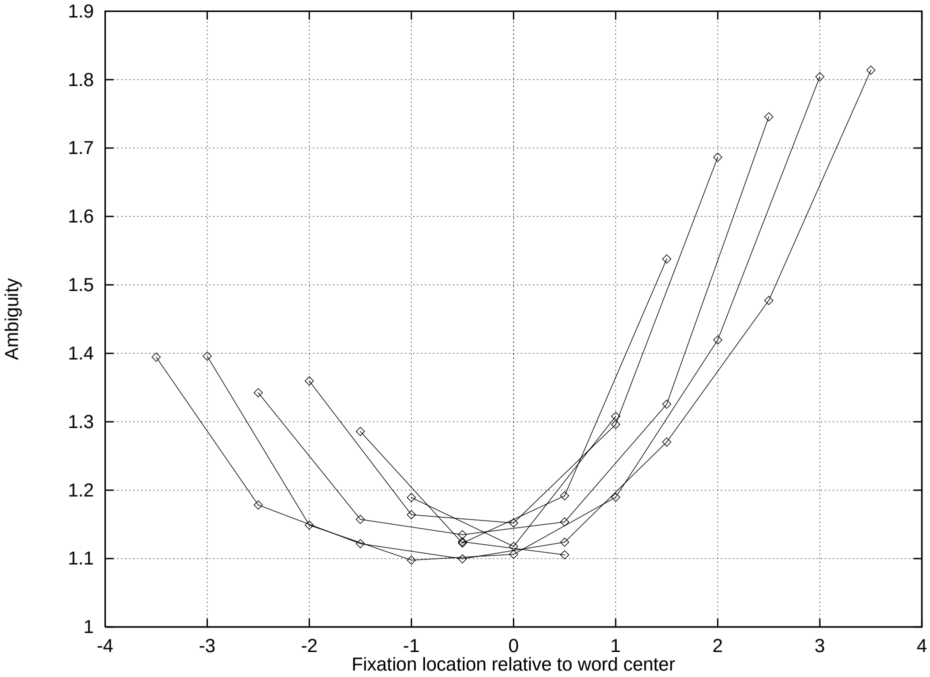

Figure 4. The ambiguity for patterns defined by the first letter, last letter and an interior letter pair, as a function of the position of the first letter in the letter pair. The different curves represent different word lengths. The Kucera and Francis word list was used to generate this graph.

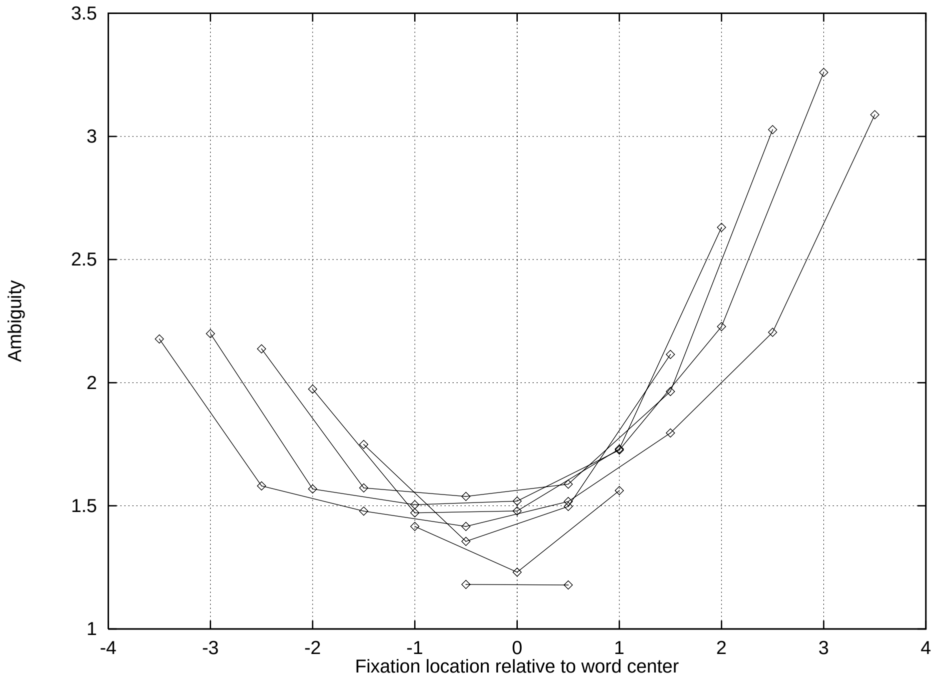

Figure 5. The ambiguity for patterns defined by the first letter, last letter and an interior letter pair, as a function of the position of the first letter in the letter pair. The different curves represent different word lengths. The Trésor de la Langue Française word list was used to generate this graph.

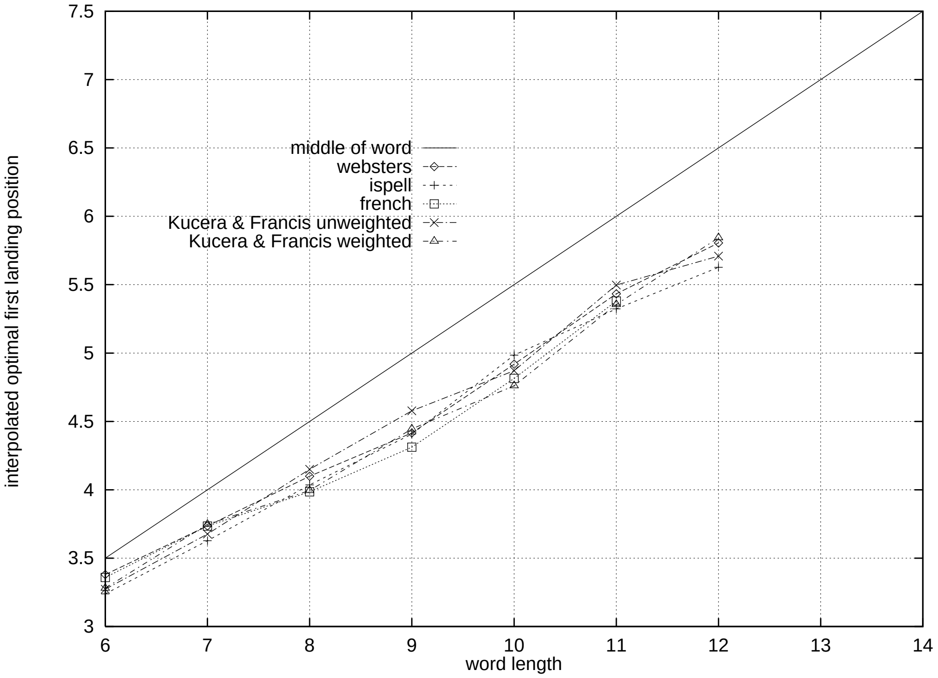

Figure 6. The (interpolated) location of the first eye fixation that yields the minimum average ambiguity in the case of recognizing interior letter pairs. The fixation location is taken to be halfway between the letters in the letter pair. Shown are the results as a function of word length for the ispell, Webster's and Kucera and Francis (with and without frequency weighting) word lists.

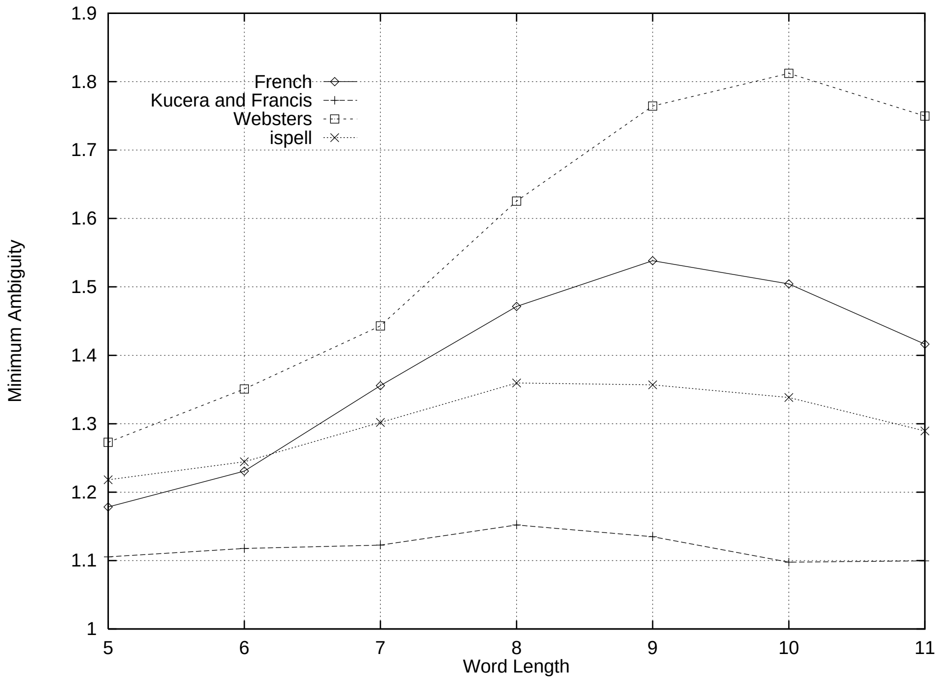

Figure 7. The value of the minimum ambiguity as a function of word length.

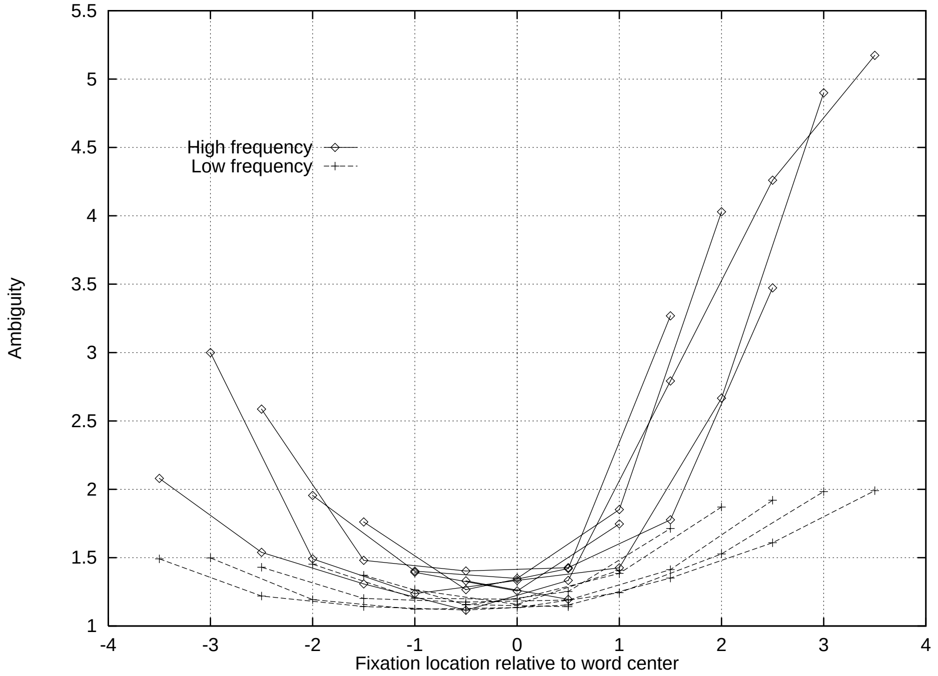

Figure 8. The ambiguity curves for the Kucera and Francis wordlist separated into high and low frequency categories.

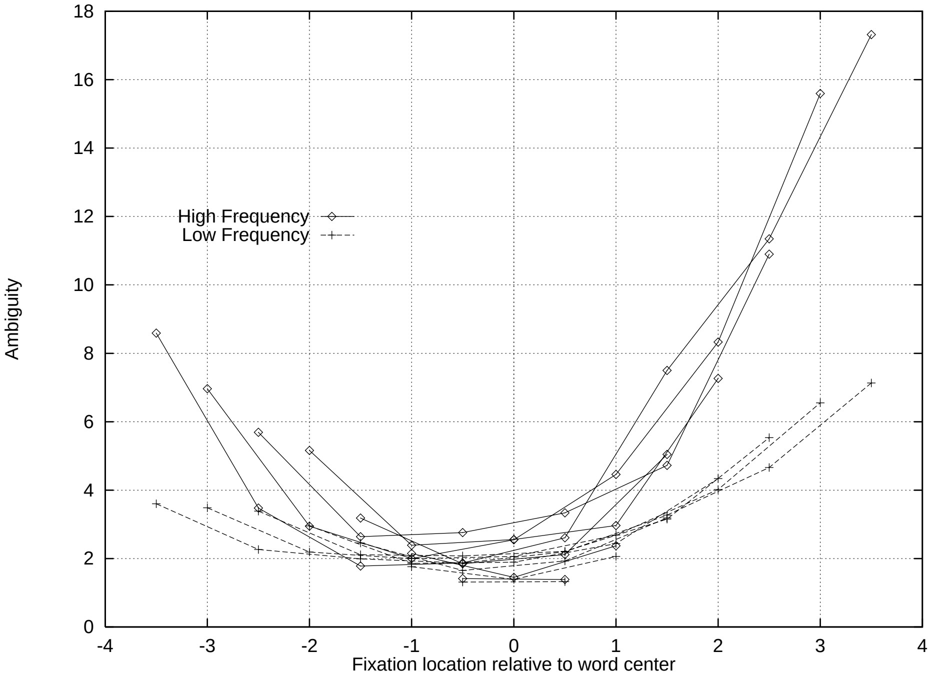

Figure 9. The ambiguity curves for the French wordlist separated into high and low frequency categories.

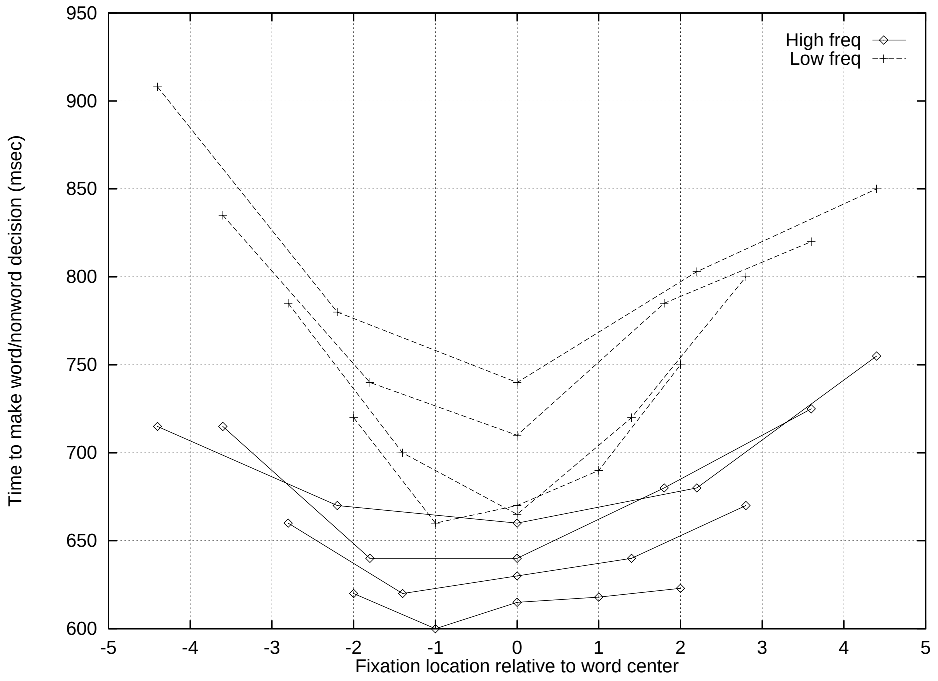

Figure 10. Data from O'Regan and Jacobs (1992) showing the time it takes to decide whether a string of characters appearing at different positions relative to the eyes' fixation point is a word or not. In the experiment, the eye fixated in the gap between two vertically aligned line segments in the middle of the screen. After 500 ms the lines disappeared and the test word or nonword appeared. The word or nonword was laterally displaced in relation to the gap in such a way that when it appeared, the eye was fixating one of five positions defined over the length of the stimulus. Curves are shown for the recognition of words of low and high frequency of occurrrence in the language, and each group of curves corresponds to words of length 5, 7, 9, and 11 letters.

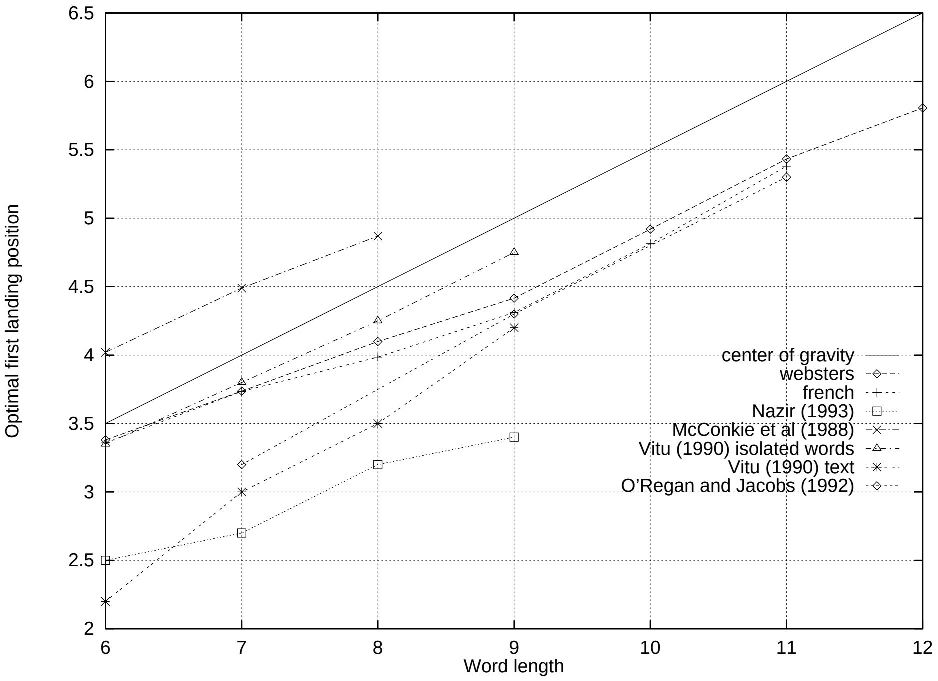

Figure 11. The interpolated position of the location where ambiguity is the smallest, or for the empirical studies, where the dependent variable (percent correct, lexical decision time, refixation probability, etc.) is at its minimum, as a function of word length.

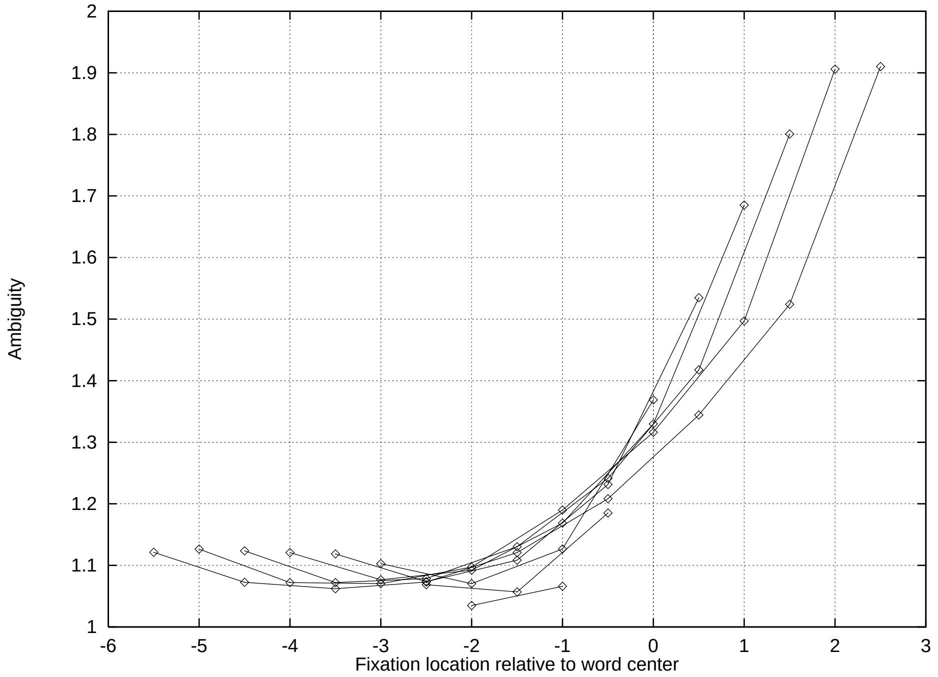

Figure 12. The minimum (over locations of the second fixation) average number of interpretations of a pattern given two fixations of the eye on the word, as a function of the location of the first fixation. At each fixation a pair of letters is recognized. In addition, the first and last letters of the word are assumed to be known. The Webster's word list was used to generate this graph.

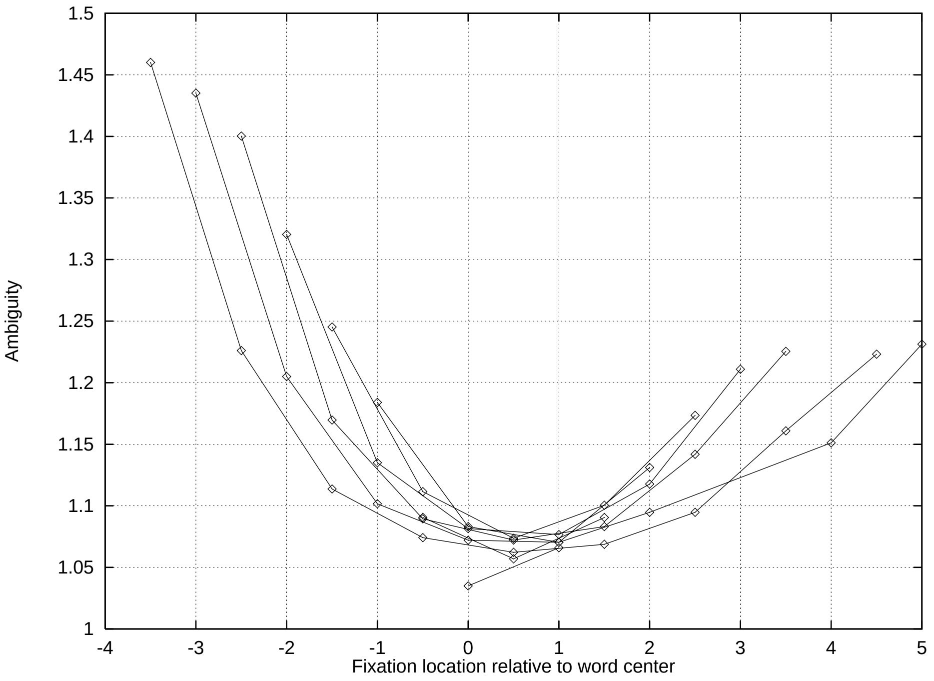

Figure 13. The minimum (over locations of the first fixation) average number of interpretations of a pattern given two fixations of the eye on the word, as a function of the location of the second fixation. At each fixation a pair of letters is recognized. In addition, the first and last letters of the word are assumed to be known. The Webster's word list was used to generate this graph.

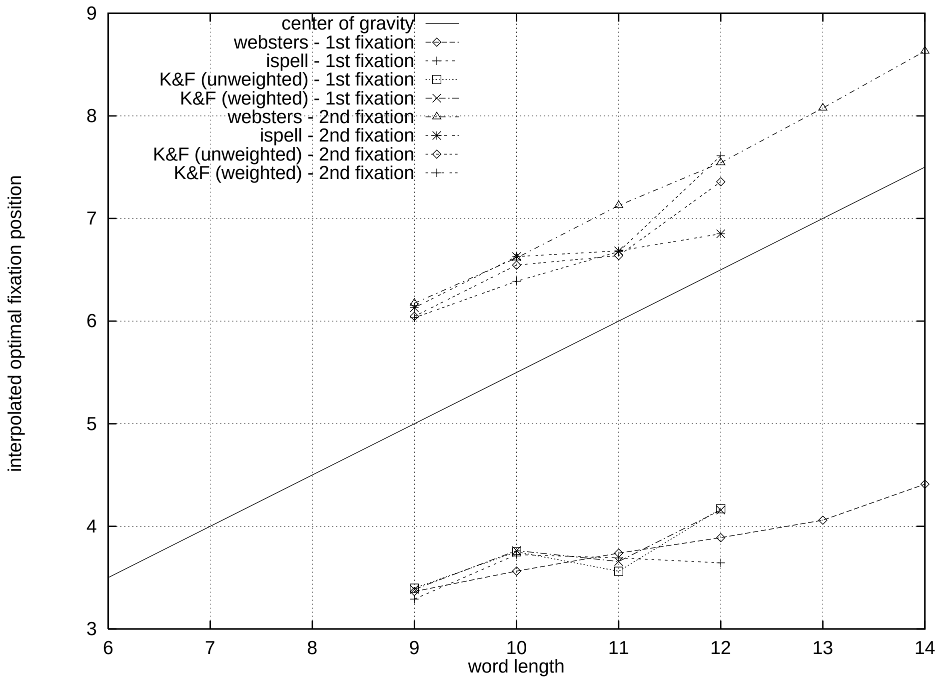

Figure 14. The (interpolated) location of the first and second eye fixations that yield the minimum average ambiguity in the case of recognizing pairs of interior letter pairs.