A Deeper Understanding of Optimal Viewing Position Using Eye Fixations and Character Recognition on Text-Viewing and Reading Tasks

Nada Attar[1], Chia-Chien[2,3], Djamel-Eddine Sia[1], and Marc Pomplun[1]

[1]University of Massachusetts, Boston, MA, USA
[2]Harvard Medical School, Boston, MA, USA
[3]Brigham and Womens Hospital, Boston, MA, USA

Abstract

Existing literature reveals that during reading, gaze fixations within a word are not necessarily close to the optimal reading position. The optimal position for reading represents the gaze fixation near the center of each word for which recognition time is minimal. In our study, we examined the match between optimal position and initial fixation position, which has not been studied before. We did this in conventional reading and in two text-viewing tasks. For the text-viewing tasks, we employed multiple words that were non-isolated in the presentation but could be examined individually in rhythm and linguistic context. We discovered that for the text-viewing tasks, the initial fixation position tends to be close to the optimal position, which was in contrast to the conventional reading task. This finding will help us understand the relationship between optimal position and reading context and inform the development of new reading applications.

Keywords: eye movements, reading, text recognition, text-viewing, optimal viewing position

Concepts: •**Human-centered computing** → *Contextual design;*

1 Introduction

During reading, we shift our line of sight to the critical positions within a text in order to recognize the words in it. The initial fixation location lies somewhere between the start and the center of the word [e.g., Reingold et al. 2012; Vitu et al. 1990], and it has been thought that readers attempt to target the centers of words, but their saccades tend to fall short [McConkie et al. 1988; Rayner 1979]. When readers eyes land at a non-optimal position in a word, they are more likely to re-fixate on the same word [ORegan 1990; Rayner and Fischer 1996] and as a result, it will take more time to recognize that word. The optimal fixation position on a word during reading, known as optimal viewing position, was discovered by O'Regan and his colleagues [O'Regan et al. 1984]. In their study, the researchers found that the probability of re-fixation on the word, as well as the fixation duration, were minimized when the initial fixation position on the word was near this point. The examination of the available literature suggests that the optimal viewing position may be asymmetrically placed, slightly left of the middle of words [ORegan et al. 1984; O'Regan and Jacobs 1992; Radach and Kempe 1993; Nazir 1993; Farid and Grainger 1996; Nazir et al. 1998]. Other studies [McConkie et al. 1988; O'Regan 1989; Nazir et al. 1991] have found the optimal viewing position to be at the middle of the word. Their hypothesis is that it might be best to place the fixation in such a way as to maximize the number of letters that are viewed with high acuity. A step towards understanding these different interpretations of optimal position is to further investigate the optimal position in different text presentation conditions. Several studies have been conducted to measure optimal viewing positions in reading isolated words. Vitu et al. [1990] designed an experiment to test whether the effect of optimal viewing position exists during the reading of words in an isolated word-reading condition, and to compare its nature and strength of the effect with that for reading of continuous text. When the initial fixation landed at the optimal viewing position, the probability of refixating within the word was much smaller than when the initial fixation landed on other positions in the word. This probability difference was rather weak in text reading but significant for isolated words. This result confirmed the existence of an optimal viewing position in both reading conditions, but its presence in texts was weaker than for isolated words, probably because of factors such as linguistic context and reading rhythm that are involved in reading texts. Moreover, the authors found that for the case of isolated words, making multiple fixations per word involves a total gaze duration penalty as compared with the single fixation case. Thus, the initial landing position in separate words must be taken into account if one wishes to understand eye-movement behavior during text reading.

Reading can also occur in other processing models than isolated words where eye movements behave differently than in conventional reading such as mindless reading [Reichle et al. 2010], scanning [Rayner and Fischer 1996], and reading words and sentences [Clifton et al. 2007]. However, there has not been as much eye-movement research in the area of text-viewing in different formats as on reading. In text-viewing, the reading process is still required as for the other models but processing the previous word is not necessary for viewing the next word. In normal reading, the fixation landing position on a word is affected by the length of the previous word and its difficulty [McConkie et al. 1988; Rayner 1979]. Text recognition does not just occur in normal reading material or reading formats but also in other circumstances, such as reading a sign during scene viewing. Therefore, text-viewing might differ from reading in terms of fixation positions, and further studies might shed light on these differences. Optimal position and fixation positions have also been investigated in object perception and in the realistic context of a natural scene. Foulsham and Kingstone [2013] explored where people fixate within photo-realistic objects and the effects of this landing position on recognition and subsequent eye movements. The results demonstrate an optimal viewing position - objects are processed more quickly when fixated at the center.

The phenomenon of optimal viewing position in words is a rich source of new ideas for understanding eye movement guidance in reading, searching, scene perception and object viewing. We introduced a new text display format for examining the optimal viewing position, as there has not been much research on such display for-

ETRA 2016, March 14 - 17, 2016, Charleston, SC, USA
ISBN: 978-1-4503-4125-7/16/03
DOI: http://dx.doi.org/10.1145/2857491.2857517

mats yet. The new presentation format that we termed text-viewing was assumed to influence eye movement behavior and add more insight into initial fixation position and optimal viewing position. Text-viewing displays contain varying numbers of words that are non-isolated in the presentation but can be examined individually in terms of rhythm and linguistic content. The words are semantically related but are isolated in their order among other words. In the present study we measure the optimal position in the traditional reading task and compare it to those obtained in two text-viewing tasks.

2 Experiment 1

We measured the position of the initial fixation on a word and its relationship to the optimal viewing position in a reading and a text-viewing task. To eliminate the effect of reading format in conventional reading, we designed a text-viewing task, where each display contained 12 words presented in random positions with a minimum distance of 5 degrees between them. Subjects examined these words individually in terms of rhythm and morphological content. We compared this task with a conventional reading task.

2.1 Experimental Setup

Eye movements were tracked and recorded using an SR Research EyeLink-2k system. Its sampling frequency was set to 1000 Hz. Stimuli (1024 x 768 pixels) were presented on a 22-inch ViewSonic LCD monitor. Its refresh rate was set to 75 Hz and its resolution was set to 1024 x 768 pixels. Participant responses were entered using a keyboard. The displays for the reading and text-viewing tasks were generated by a MATLAB script. The average length of words for both tasks was 6.4 letters (minimum = 3 letters, maximum = 13 letters). The reading task included words of length 1 and 2 but they were excluded from the analysis because the text-viewing task did not have these two lengths. We used Courier font and font size 20pt. The font color was black and the background color was grey.

To avoid learning effects or any influence of one of the tasks on the other, we used different sets of words when creating the stimuli for the two tasks. The passage for the reading contained general topics (food, science, health, history) taken from Washington Post news. The word categories for text-viewing were taken from 30 different categories (e.g. airport, beach, bedroom, or library). There were six blocks in the reading task. Every block included one passage divided into 3 parts to give a total of 18 displays to the users. Each passage had between 230 and 240 words. The total of 42 stimuli for the text-viewing task was divided into six blocks, with every block including seven trials. The words in every trial belonged to one, two or three categories. For example: words such as airplane, baggage, and boarding belonged to one category “airport”whereas words such as blanket, mattress, and surfboard belonged to two categories “bedroom”and “beach”.

2.2 Procedure

Twenty one subjects, native English speakers, aged between 18-32 years old, were tested. All had normal or corrected to normal vision and were nave to the purpose of the study. An instruction screen was shown before each block to inform participants about the type of the task. The order of the tasks was counterbalanced. Before the actual experiment, subjects had one block of practice in each task. Subjects were trained to read and not to scan in the reading blocks, and to examine most of the items on the text-viewing display before they identified the number of categories and input the number of categories via the keyboard.

In the reading task, each block had one passage divided into three screens. Subjects were trained to read the text on the first screen, then press the space bar to move to the second screen, and press it once more to view the third screen. The screen would automatically move to the next one after a 60-second timeout if they did not press the space bar. When subjects pressed the space bar after reading the third screen, a blank screen appeared. The experimenter would ask comprehensive questions from the passage and subjects had to verbally report their answers. The experimenter would instruct them when to start the next trial. During each trial in the text-viewing task, subjects were instructed to identify the number of categories to which the 12 words belonged and respond by pressing the numbers (1, 2 or 3) on the keypad. The next trial would begin once subjects pressed any of those keys or the stimulus had been shown for 50 seconds (timeout). After subjects responded, a feedback sounds was played to indicate whether the subjects answer was correct or incorrect before a new trial appeared.

2.3 Data Analysis

For the reading and text-viewing tasks, we only took in consideration the initial fixation that landed on each word and ignored the next fixations as well as the revisiting fixations for that particular word. Since all the words in text-viewing task were nouns, we only considered the nouns in the analysis of the reading task. This left us with 450 nouns in the reading task compared to 528 words in the text-viewing task after eliminating the second part of the combination words in the text viewing task. The difference in the number of words between the tasks was not significant for any word length, $t(12) = 1.24$, $p > .05$. Our fixation analysis method collected the letters that were targeted by the subjects fixations. We used a MATLAB function from the Computer Vision System Toolbox to perform optical character recognition (OCR) on each display individually (reading display or text-viewing stimulus) by extracting all the words with their properties and characters. OCR allows experimenters to turn any reading material into stimuli for experiments or enhance its readability based on experimental results without the need to define areas of interest in advance. Every word received a bounding box with a margin of 1° at its top and bottom and horizontal margins equally filling the gaps between consecutive words. Every letter had its own bounding box as well.

2.4 Results

The question in the current study was whether readers fixation locations on the words were different when they read a passage than when they read words in the text-viewing task. We were also interested in whether the initial fixation position during text-viewing was close to the optimal viewing position. For each word length by each subject, we assigned every fixation to the letter in whose bounding box it landed. Then the mean letter position was calculated for the subjects fixations on all words of a given length, and these means were averaged across all subjects. We found that in the reading task, the mean fixation position results were similar to the findings of prior studies in that the initial fixations landed between the beginning and the middle of a word. The mean initial fixation position for the text-viewing was located at the middle of the word. The average of each subjects mean of the fixation position for reading was at the 2.7th letter (SD = 0.73) for the average word of 6.33 letters. In the text-viewing task, the average of each subjects mean of the fixation landing site was at position 3.61 (SD = 0.42) for the average word of length 6.4. The difference in fixation position between the two tasks across subjects was statistically significant, $t(20) = 2.22$, $p < .005$.

In order to further investigate how the fixation position varied as a function of the words length, we examined the initial fixation landing position for every word length from three to 13 for both tasks. For each word length, the mean of each subjects results related to the words of that length was calculated and the means were averaged across subjects. Figure 1 shows how in the reading task the fixation position deviated further from the center of the word when the word length increased. In contrast, the initial fixation position in the text-viewing task tended to remain closer to the center of the word. Figure 1 shows the mean initial fixation landing position for every word length in each of the two tasks.

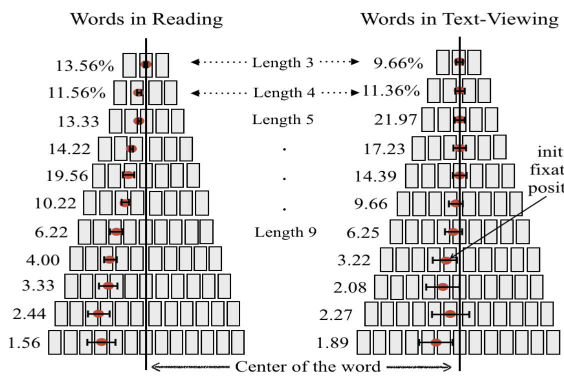

Figure 1: *Initial fixation positions across word lengths in each of the two tasks. The red dots show the average of subject's mean of the initial landing positions in each task for all subjects. Horizontal bars indicate the standard error of the mean, and numbers indicate word length frequencies.*

2.5 Discussion

The goal of the first experiment was to investigate whether the initial fixation positions in the reading task differed from those in the text-viewing task. We did find this difference; however, it is important to notice that these values of fixation position are averages, and there is considerable variability in both tasks. Particularly, based on the frequencies of the words (see Figure 1), the data of the long words are insufficient for drawing strong conclusions and will require further studies. The left bias found in the reading task can be strongly influenced by the format of presentation, rather than by the nature of the reading behavior. In reading, fixations are typically made after a saccade that has a direction from the left to the right (or right to left if the word is revisited). Conversely, the results we obtained in the text-viewing task were after saccades that could come from any direction or angle. Therefore, there is a need for an additional experiment in which the presentation of the words in the text-viewing task can be examined in the same way as in the conventional reading, that is, in a top-to-bottom and left-to-right manner. We examined this possibility in Experiment 2.

3 Experiment 2

We presented an additional comparison between the initial fixation position in reading task and in a modified text-viewing task. The frequency of the word length in this task was also not significantly different from the reading task, $t(10) = 1.76$, $p > .05$.

3.1 Experimental Setup

The apparatus had the same setup as Experiment 1. The displays for the text-viewing task were again generated by a MATLAB script to avoid stimuli being identical to Experiment 1 with regard to words and their categories. The average length of words for the text-viewing task was six letters (minimum = 3 letters, maximum = 13 letters). We used same font, background design, and same 30 categories for the text-viewing task as in Experiment 1. The text-viewing task had 42 stimuli divided into six blocks, and every block consisted of seven trials. Every trial had a total of 12 words that belonged to either one, two or three categories. Subjects performed an additional block of practice trials prior to the experiment. Each stimulus display was composed of three categories. The words were aligned line by line with 3 words per line and a minimum distance of 5 degrees between the words.

3.2 Procedure

The same 21 subjects from Experiment 1 participated again in Experiment 2, with a short break between experiments. The procedure and the flow of the experiment were identical to the text-viewing condition in Experiment 1 but participants were trained to look at the 12 words in the display in the same order in which they read, line by line starting from the top left and ending at the bottom right as they were trained to do in the practice block (see Figure 2).

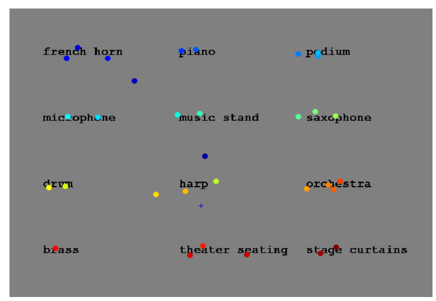

Figure 2: *A plot of one subjects initial fixations in Experiment 2. The dots are the fixations and were colored sequentially based on their times and the reading direction.*

3.3 Data Analysis

Same as in Experiment 1.

3.4 Results

In the same way as in Experiment 1, every fixation was assigned to the letter in a word in whose bounding box it landed. Subsequently, the mean was calculated for all subjects fixations on that word. The average of each subjects mean of the fixation landing site was letter 3.23 ($s = 0.41$ letters) for the average words of six letters. The difference between subjects fixation positions in the text-viewing task in Experiment 2 and the reading experiment in Experiment 1 was statistically significant, $t(20) = 1.90$, $p < .05$. The difference between subjects fixation positions in the text-viewing tasks of the experiments was not statistically significant, $t(20) = 1.6$, $p > .05$. To further investigate how the fixation position varied as a function of the words length, we examined the initial fixation position for every word length from three to 13. The mean of all subjects results related to each word length was calculated. Figure 3 shows that the fixation position in the text-viewing task was near the center of the word, similar to the results of the text-viewing task in Experiment 1.

3.5 Discussion

The purpose of Experiment 2 was to measure the initial fixation positions for a text-viewing task with minimal difference in presentation to the reading task. Interestingly, the results showed that the type of presentation did not affect the initial fixation position when words were being read separately. Examination of the initial fixation position on the displayed words demonstrated that viewers viewing location is close to the optimal viewing position in both text-viewing tasks of Experiments 1 and 2, and the leftward bias occurs only in the reading behavior. Therefore, it seems that these landing positions are actually task dependent and are close to the center position when there is no actual reading behavior. The left bias comes purely from the nature of reading, not the format of reading. However, more data points are required to strengthen the results for long words.

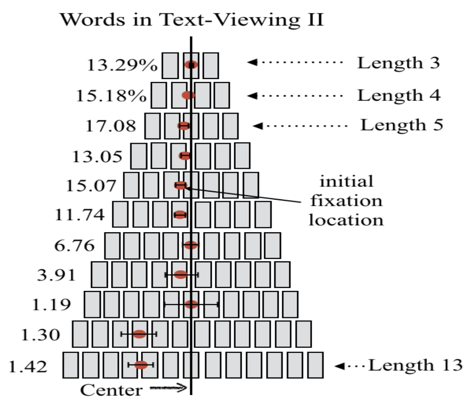

Figure 3: *Initial fixation positions as a function of word length in Experiment 2. The red dots show the average of subjects means of the initial landing positions. Horizontal bars indicate the standard error of the mean, and numbers indicate word length frequencies.*

4 Conclusion

Prior studies have shown that, for conventional reading tasks, the fixation position tends to be left of the optimal viewing position. In our study, we investigated initial fixations in text-viewing tasks and discovered that the fixation position tends to be closer to the optimal viewing position. Our second experiment confirmed that the initial fixation landing position results were not biased by the saccades that can come from different directions during viewing. This work helps us understand the match between optimal viewing position and initial fixation position in different reading contexts in order to design better reading applications. More work is still required on the influence of word semantics and familiarity in text-viewing tasks.

References

CLIFTON, C., STAUB, A., AND RAYNER, K. 2007. Eye movements in reading words and sentences. *Eye movements: A window on mind and brain*, 341–372.

FARID, M., AND GRAINGER, J. 1966. How initial fixation position influences visual word recognition: A comparison of french and arabic. *Brain and Language 53*, 3, 351–368.

FOULSHAM, T., AND KINGSTONE, A. A. 2013. Optimal and preferred eye landing positions in objects and scenes. *The Quarterly Journal of Experimental Psychology 66*, 1–22.

MCCONKIE, G. W., KERR, P. W., REDDIX, M. D., AND ZOLA, D. 1988. Eye movement control during reading: I. the location of initial eye fixations on words. *Vision Research 28*, 1107–1118.

NAZIR, T. A., OREGAN, J. K., AND JACOBS, A. M. 1991. On words and their letters. *Bulletin of the Psychonomic Society 29*, 171–174.

NAZIR, T. A., JACOBS, A. M., AND O'REGAN, J. K. 1998. Letter legibility and visual word recognition. *Memory and Cognition 26*, 4, 810–821.

NAZIR, T. A. 1993. On the relation between the optimal and the preferred viewing position in words during reading. *In: G dYdewalle, J van Rensbergen (Eds.). Perception and cognition: advances in eye movement research*, 349–361.

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 18*, 185–197.

O'REGAN, J. K. 1984. How the eye scans isolated words. *In Gale, A. G. Johnson, F. (Eds.), Theoretical and Applied Aspects of Eye Movement Research*, 159–168.

O'REGAN, J. K. 1989. Visual acuity, lexical structure, and eye movements in word recognition. *In: Elsendoorn, B, and Bouma, H (Eds.), Working Models of Human Perception*, 261–292.

O'REGAN, J. K. 1990. Eye movements and reading. *In E. Kowler (Ed.), Eye movements and their role in visual and cognitive processes*, 395–453.

OREGAN, J. K., LE'VY-SCHOEN, A., PYNTE, J., AND BRUGAILLE'RE, B. 1984. Convenient fixation location within isolated words of different length and structure. *Journal of Experimental Psychology: Human Perception and Performance 10*, 250–257.

RADACH, R., AND A KEMPE, A. 1993. An individual analysis of initial fixation position in reading. *In: G dYdewalle, J van Rensbergen (Eds.), Perception and cognition: advances in eye movement research*, 213–225.

RAYNER, K., AND FISCHER, M. H. 1996. Mindless reading revisited: Eye movements during reading and scanning are different. *Perception and Psychophysic 58*, 734–747.

RAYNER, K. 1979. Eye guidance in reading: Fixation locations in words. *Perception 8*, 21–30.

REICHLE, E. D., REINEBERG, A. E., AND SCHOOLER, J. W. 2010. Eye movements during mindless reading. *Psychological Science 21*, 9, 1300–1310.

REICHLE, D., E., GLAHOLT, M. G., AND SHERIDAN, H. 2012. Direct lexical control of eye movements in reading: Evidence from survival analysis of fixation durations. *Cognitive Psychology 65*, 177–206.

VITU, F., OREGAN, J. F., AND MITTAU, M. 1990. Optimal landing position in reading isolated words and continuous text. *Perception and Psychophysics 47*, 583–600.