Attentional Processes in Natural Reading: the Effect of Margin Annotations on Reading Behaviour and Comprehension

Andrea Mazzei* Tabea Koll† Frédéric Kaplan‡ Pierre Dillenbourg§
School of Computer and Communication Sciences, École Polytechnique Fédérale de Lausanne
Albert-Ludwigs-Universität Freiburg

Abstract

We present an eye tracking study to investigate how natural reading behavior and reading comprehension are influenced by in-context annotations. In a lab experiment, three groups of participants were asked to read a text and answer comprehension questions: a control group without taking annotations, a second group reading and taking annotations, and a third group reading a peer-annotated version of the same text. A self-made head-mounted eye tracking system was specifically designed for this experiment, in order to study how learners read and quickly re-read annotated paper texts, in low constrained experimental conditions. In the analysis, we measured the phenomenon of annotation-induced overt attention shifts in reading, and found that: (1) the reader's attention shifts toward a margin annotation more often when the annotation lies in the early peripheral vision, and (2) the number of attention shifts, between two different types of information units, is positively related to comprehension performance in quick re-reading. These results can be translated into potential criteria for knowledge assessment systems.

CR Categories: H.5.2 [Information Interfaces and Presentation]: User Interfaces—Input devices and strategies;

Keywords: Eye movements, Reading Comprehension

1 Introduction and Motivation

Nowadays students use personal annotation strategies on printed instructional materials in their daily study activities. They highlight; they write short notes within the margins or between lines of text; they use longer notes and graphical sketches in blank spaces or near figures. A vast area of educational research has associated note-taking and note-reading to enhanced learning outcomes. In addition, the presence of extraneous annotations on instructional texts can support learners with low prior knowledge to better understand the content. In this work we study how people read and reread in the presence of annotations, spontaneously taken during a natural reading session, and how the attentional processes involved relate to the learning outcome. Our experimental investigation aimed at emulating the natural reading and annotating experience, by recreating the conditions characterizing the students' study activity and by mitigating the effects of the experimental constraints on the task performance. To study the learners' reading behavior, a head mounted eye tracking system was employed in the experiment.

*andrea.mazzei@epfl.ch
†tabeakoll@hotmail.com
‡frederic.kaplan@epfl.ch
§pierre.dillenbourg@epfl.ch

2 Related work and Research Questions

2.1 Visual Saliency and Eye Movements in Reading

The literature of eye movements in reading is a vast area of research that is attracting practitioners from different disciplines. The survey paper published by [Rayner 2009] gives detailed overviews on studies focusing on eye movements and attention in reading. Many of the empirical studies reported in Rayner's paper discuss how readers' attention is influenced by low-level visual information. Some of the reported effects were measured through fixations and forward saccadic distributions. For instance, the fixation duration, the probability of fixating on words and the saccadic distribution are influenced by topological factors of the text [Juhasz et al. 2008]. Some eye tracking studies have investigated how viewers integrate pictorial and textual information. [Rayner et al. 2001] showed that the size of the print has an effect on the fixation duration and saccadic length. [Pieters and Wedel 2004] elaborated an attentional model in mixed textual and pictorial stimuli, and used it to study attentional processes in brand advertising. They found evidence for an effect of text and pictorial size on incremental attention capture and transfer. Based on the literature discussed here we can infer that low level visual information has an influence on the saccadic distribution in reading. Therefore we here formulate our first research question applied to the domain of natural reading.

$\mathbf{RQ_A}$. How does the presence of annotations affect reading?

Our hypothesis of attention transfer in natural reading takes into account only margin annotations. We did not consider highlighting and underlining, because they were reported to have weak implications on readers' recall and learning.

- RH_1. *Margin annotations attract the readers' attention proportionally to measures of proximity in the visual field.*

2.2 Visual Attention and Learning

More controversial and non-replicated results have related semantic processing with eye movement control. If the *where* in reading is influenced by low level visual cues, the *when* seems to be determined by cognitive processes [Pereaa et al. 2003]. [Schmidt-Weigand et al. 2010] found that visual attention in learning tasks is predominantly driven by the textual components. They conducted two experiments focusing on visual attention in learning from mixed text and pictorial stimuli. They observed that, in the written text modality, the participants alternated between text and pictorial elements frequently. No conclusions on learning performance could be drawn from the two analyses. [Cromley et al. 2010] found that, in a mixed text and diagram stimuli, students tend to avoid diagrams, even though these summarize the textual content. The authors also found that, the time spent reading the text vs. the time spent studying the diagrams, was not related to the learning outcome. On one side, these two studies failed at measuring the relationship between readers' attention shifts from textual instructions towards pictorial complements and learning outcome. On the other side, the pedagogical research on note-taking and note-reading is generally supportive of the beneficial effects on reading comprehension and recall. Therefore, we decided to investigate potential relationships between annotation-induced attention shifts and reading comprehension. To measure the shift of attention we studied gaze-shift dynamics. More precisely, in our context, we define gaze shift as a redirection of the reader's attention towards a region of the text containing an information item different than the one observed. Hence, we formulated our second research question as follows:

$\mathbf{RQ_B}$. Are the annotation-induced gaze shifts related to the reader's comprehension?

Our expectation is that, up to a certain frequency, these gaze shifts are proportional to the reader's comprehension performance.

- RH_2. *The rate of gaze shifts induced by annotations while reading the instructional text is positively related to the reader's comprehension.*

Some other studies have elaborated cognitive processes related to backward and global regressions. For instance, regressive eye movements predominantly relate to previously read information units that are meaningful for the reader [Kennedy et al. 2003]. The spatial encoding theory on regressive saccades was empirically observed by many authors. Readers remember where the cognitive operation took place [Baccino and Pynte 1994; Kennedy 1992]. More theories strengthen the relationship between spatial encoding processes and text comprehension [Fischer 1999]. [Inhoff and Weger 2005] discussed the spatial indexing hypothesis: readers memorize the position of identified information units. This last section of the literature review suggested us to introduce a quick rereading phase into our experiment design. In order to investigate these effects of annotations in rereading we formulated our third hypothesis.

- RH_3. *The rate of gaze shifts induced by annotations while quickly rereading the instructional text is positively related to the reader's comprehension.*

3 Method

A between-subjects experimental design was chosen to investigate the effect of note-taking and reading peer-annotated documents on learning. The experiment was characterized by 3 conditions and 2 experimental phases. In the first experimental phase (*reading*) the subjects had enough time to read and understand the instructional material. In the second experimental phase (*rereading*) the subjects came back after one week to quickly reread the same text. Three experimental groups were designed for this experiment. The participants in the *read_only* group read and reread the unannotated text. This group served as control for the other two experimental groups. The participants in the *self_annotation* group read the instructional text and, while reading it, they were allowed to take annotations. The participants in the *peer_annotation* group read the text that was chosen as the best annotated one from the pool of participants in the *self_annotation* group during the *reading* phase.

3.1 Participants

A total of 56 native French-speaking first and second year university students participated in the experiment. The subjects were assigned to the 3 conditions as follows: *read_only*: N = 16 (11 males, 5 females), *self_annotation*: N = 24 (17 males, 7 females), *peer_annotation*: N = 16 (10 males, 6 females).

3.2 Instructional Materials

The instructional text was developed by an expert in the field and written in French. It consists of six pages, 1546 words, explaining the dynamics underlying the phenomenon of neural signal transmission. The text does not contain any illustration, tables, or schemas. The subjects in the *peer_annotation* group read an instructional text which had been previously annotated by a subject in the *self_annotation* group. Among all annotated documents produced by the subjects in the *self_annotation* group, the best one was chosen as follows. First, only the six students who achieved the highest 1st post test score were taken into consideration. Second, these 6 annotated documents were ranked based on four different criteria: the balance among the different annotations strategies, the readability of the annotations, the semantic correctness, and the semantic coverage of the information units.

3.3 Apparatus

During the experiment, each subject was asked to wear our dual camera head-mounted eye tracking system [Mazzei et al. 2014]. The eye tracker permits one to study natural reading in low constrained experimental settings, and allows for free head movement and note-taking. The system is capable of extracting the subject's gazepoint on paper semi-automatically, offline, and without requiring online calibration procedures. Gazepoint measurements are post-extracted at a 75 HZ rate. We conducted an evaluation study of the system and we measured an average precision of 0.87 degrees and an accuracy of 1.01 degrees.

3.4 Knowledge Assessment Tests

Prior knowledge is an essential variable in the process of knowledge assimilation. Therefore, we tailored a pre-test to be used to assess prior knowledge and to statistically control for it in the analyses. It was developed by two of the authors, and consisted of 24 true or false questions (8 questions for each of the 3 sections of the text) covering the most salient aspects. Our two post-tests consist of 48 questions, 16 questions per section of the instructional text. The 48 questions are divided in 24 multiple choice questions, and 24 inference questions (8 questions per text section for both types).

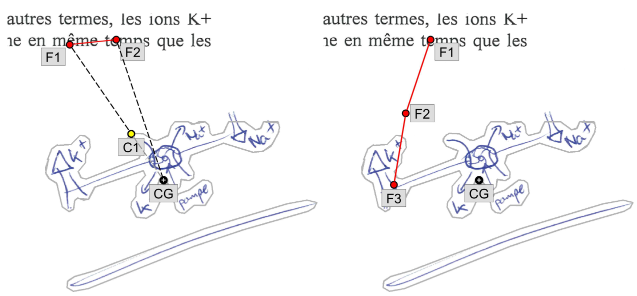
4 Data Collection & Processing

4.1 Visual Features of the Information Items

The annotations, produced by the subjects in the *self_annotation* group, were manually identified based on semantic completeness. Their contour and center of gravity (CG) were automatically computed using a computer vision procedure embedded in our framework [Mazzei et al. 2014]. The instructional material and the posttests were structured in 3 independent sections, to perform a repeated measures analysis. In addition, all the gaze features computed for the analysis were labeled in order to specify the spatial relationship with the type of information item, or a combination of them: paragraph (*P*), annotation (*A*).

4.2 Saccadic Length vs Annotation Proximity (*SLAP*)

In Section 2.1, we hypothesized that annotations exert an attractional influence on the generation of the saccadic flow while reading, based on a principle of spatial proximity. Inspired by the gaze analyses reported in Section 2.1, we defined SLAP (see Figure 1(a)), a gaze reading feature that contrasts the saccadic distribution with the spatial proximity of the closest annotation. Figure 1(a) contains a red segment, $\overline{F_1F_2}$, corresponding to a reading saccade. The saccade is delimited by two fixations F_1 and F_2. Given the position of fixation F_1, the closest among all the annotations present on the same page is identified. Finally, the segment $\overline{F_1CG}$ and the



(a) *SLAP* (b) *PA_gaze_shift*

Figure 1: *Gaze features over a portion of an annotated page*

resulting intersection with the contour C_1, were found. The length of the segment $\overline{F_1C_1}$ is the x coordinate of the *SLAP* feature. The length of the segment $\overline{F_2CG}$ is the y coordinate. For this second distance we considered the annotation CG instead of the contour because some of the saccades land inside the contour.

4.3 Gaze Shift

In our experimental context, we do not discriminate between the first gaze shift and regressive movements toward the same location. We identified the gaze shift as a chain of sequential long saccades linking the two regions of interest. The figure 1(b) shows a type of gaze shift often used in this analysis: the *PA_gaze_shift*. In this example, the link connects the word of the text *"temps"* with the left part of the corresponding annotation.

5 Results

In this section we present the results of regressional analyses in order to test our research hypotheses. In these analyses we did not test for differential relationships between gaze shifts and learning outcomes in the three conditions.

5.1 Saccadic Processing and Annotation Proximity

The scatter plot in Figure 2 visualizes the *SLAP* gaze feature for all the saccades recorded from all the subjects in the *self_annotation* condition during the *rereading* phase and *peer_annotation* condition in both experimental phases. The data from the *self_annotation* condition in the *reading* phase were excluded because they involve not only reading but also annotating. Each point in the plot corresponds to a saccade and has two coordinates:

- **x coordinate** is the distance between the starting fixation (F_1) and the contour of the closest annotation (C_1),
- **y coordinate** is the distance between the ending fixation (F_2) and the CG of the annotation.

The plot reveals two main regions. The first main central region of points lies around the first quadrant bisector. These points represent those saccades that kept a similar distance between the two fixations position and the closest annotation. Therefore we could infer that they were not influenced by the proximity of the closest annotation. Below the first quadrant bisector there is a second smaller cluster of saccades. These saccades start far from the closest annotation contour but land close to the CG of the same annotation. This suggests that these saccades are subject to an attraction towards the closest annotation. These observations justified a quantitative investigation to obtain better understanding of the nature of this phenomenon.

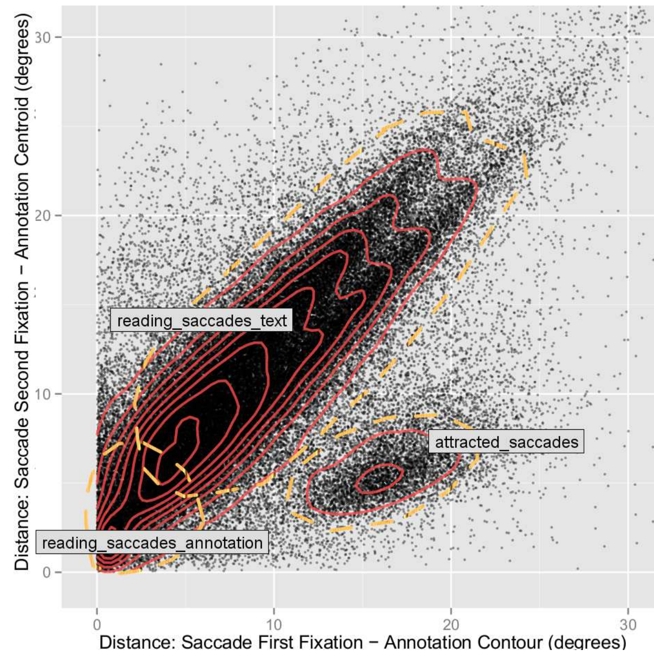


Figure 2: SLAP *gaze feature points*

5.2 Gaze Shifts from Paragraph to Annotations

In this section we address the research hypothesis RH_1 by studying the *PA_gaze_shift*. Figure 3 displays the histograms of the number of *PA_gaze_shift*, sorted by length; a color code is used to differentiate the regions of visual acuity of the human eye [Poppel and Harvey 1973]. The histograms, shown in Figure 3, suggest that there is a predominant role of the early peripheral vision, *plateau*, in the generation of *PA_gaze_shift*. To investigate this we restructured the data set containing the *PA_gaze_shift* entries. We categorized the length of each gaze shift entry based on the related region of visual acuity. We built a linear mixed-effect model to relate the regions of visual acuity (categorical predictor) with the count of *PA_gaze_shift* (criterion variable). The resulting test was significant ($F(3, 1125) = 229.22$, $p < .01$). The result indicates that there is a significant difference in the rate of *PA_gaze_shift* among the different regions of visual acuity. From a graphical analysis of the histograms in Figure 3, we can confirm this result. The spatial proximity of an annotation affects the reading behavior by inducing shifts of attention. This happens with highest frequency when the annotation is in the *plateau* of the reader's vision. This analysis presents a limitation. Margin annotations are, by definition, distant from the body of the text. Therefore, they tend to be less linked when in the foveal or perifoveal regions. On the other hand, the degrading trend after the first peak at 8 degrees suggests that the farther the point of gaze is from the closest annotation, the less frequently the reader's attention is attracted by the annotation. Therefore, we accept RH_1.

5.3 Gaze Shifts and Learning Outcome

In this Section we address the research hypotheses RH_2 and RH_3. We here investigate whether gaze shifts are related to the reader's comprehension. In the analysis we employ linear regression models because the sample size is limited and because, at this stage, we only intend to point out macroscopic trends. We do not exclude that a similar analysis with a higher sample size could be addressed using logistic regression. A series of linear mixed-effects models

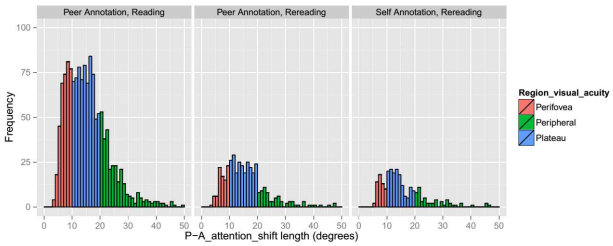


Figure 3: *Histograms of the* PA_gaze_shift

were built in order to assess whether there is a significant relationship between the number of various gaze shifts and the two post-test scores. Considering that each user had three measures, the model took into account the user identifier as grouping factor. The pretest score was used as covariate. The results concerning the first *reading* phase and only the *peer_annotation* condition are as follows.

- The *PP_gaze_shift* predicts ($Est. = 0.02$, $Std.Err. = 0.006$) the 1st post test score ($F(1, 30) = 5.30$, $p = .03$).
- The *PA_gaze_shift* does not predict the 1st post test score ($F(1, 30) = 0.01$, $p > .05$).
- The *AA_gaze_shift* does not predict the 1st post test score ($F(1, 30) = 0.02$, $p > .05$).

All the models resulted to be not significant except for the one having the *PP_gaze_shift* as predictor in which the effect is modest. Therefore we can conclude that in the *reading* phase and *peer_annotation* condition the number of gaze shifts does not predict the reading comprehension, and we reject RH_2. The results concerning this second experimental phase and both the *self_annotation* and *peer_annotation* are as follows.

- The *PP_gaze_shift* predicts ($Est. = 0.15$, $Std.Err. = 0.072$) the 2nd post test score ($F(1, 59) = 4.54$, $p = .03$).
- The *PA_gaze_shift* predicts ($Est. = 0.11$, $Std.Err. = 0.036$) the 2nd post test score ($F(1, 59) = 8.64$, $p < .01$).
- The *AA_gaze_shift* predicts ($Est. = 0.07$, $Std.Err. = 0.027$) the 2nd post test score ($F(1, 59 = 7.62$, $p < .01$).

All tests showed a significant positive relationship. Therefore we can conclude that the number of gaze shifts between two paragraphs (*PP*), between a paragraph and annotation (*PA*) and between annotation and annotation (*AA*), are potential positive predictors of the fast re-reading comprehension, and we accept RH_3.

6 Discussion

In this study, we have provided a gaze-centered perspective on how student-annotated instructional texts are read. Margin annotations attract the reader's attention based on spatial proximity. This result is compatible with the models of attention transfer in mixed textual and pictorial stimuli discussed by [Rayner et al. 2001; Pieters and Wedel 2004]. In addition we have related some of the gaze features to reading comprehension performance. Gaze shifts, between different types of information items, resulted to be positively related to the comprehension score in fast re-reading. This result suggests that gaze shifts constitute a feature that can be used to explain a part of the reading comprehension process. This result is a complement to existing empirical observations: regressive eye movements in reading are a manifestation of intelligent spatial control of reader's attention [Kennedy et al. 2003]. Our interpretation is that annotations in a re-reading task constitute a manifestation of the spatial indexing hypothesis [Inhoff and Weger 2005]: students go back to the elements they remembered as being important when reading a text. Our gaze analysis makes us reflect on the potential of eye tracking methodology in supporting knowledge assessment systems, where the use of tests is not practicable. Due to the rather small sample size and only one experimental setup, generalizability of our results is limited. However, from the obtained data, it is already possible to further elaborate gaze-and-annotation-based models capable of predicting reading comprehension. Our results should inspire other researchers to further explore this promising line of research.

References

BACCINO, T., AND PYNTE, J. 1994. Spatial coding and discourse models during text reading. *LCP 9*, 143–155.

CROMLEY, J. G., SNYDER-HOGAN, L. E., AND LUCIW-DUBAS, U. A. 2010. Cognitive activities in complex science text and diagrams. *CEP 35*, 59–74.

FISCHER, M. H. 1999. Memory for word locations in reading. *Memory 7*, 79–116.

INHOFF, A. W., AND WEGER, U. W. 2005. Memory for word location during reading: Eye movements to previously read words are spatially selective but not precise. *Memory & cognition 33*, 3, 447–461.

JUHASZ, B., WHITE, S., LIVERSEDGE, S., AND RAYNER, K. 2008. Eye movements and the use of parafoveal word length information in reading. *JEPHPP 34*, 6, 1560–1579.

KENNEDY, A., BROOKS, R., FLYNN, L. A., AND PROPHET, C. 2003. The reader's spatial code. In *European conference on eye movement research; The mind's eye 11th, European conference on eye movement research; The mind's eye; 193-212*, 193–212.

KENNEDY, A. 1992. The spatial coding hypothesis. In *Eye movements and visual cognition*. Springer, 379–396.

MAZZEI, A., EIVAZI, S., MARKO, Y., KAPLAN, F., AND DILLENBOURG, P. 2014. 3d model-based gaze estimation in natural reading: a systematic error correction procedure based on annotated texts. In *Proceedings of the Symposium on Eye Tracking Research and Applications*, ACM.

PEREAA, M., CARREIRASB, M., PEREAA, M., AND CARREIRASB, M. 2003. Regressions and eye movements: Where and when. *BBSU 26*, 4.

PIETERS, R., AND WEDEL, M. 2004. Attention capture and transfer in advertising: Brand, pictorial, and text-size effects. *Journal of Marketing*, 36–50.

POPPEL, E., AND HARVEY, L. O. 1973. Light-difference threshold and subjective brightness in the periphery of the visual field. *Psychologische Forschung 36*, 2, 145–161.

RAYNER, K., ROTELLO, C., STEWART, A., KEIR, J., AND DUFFY, S. 2001. Integrating text and pictorial information: Eye movements when looking at print advertisements. *JEPA 7*, 3, 219.

RAYNER, K. 2009. Eye movements and attention in reading, scene perception, and visual search. *QJEP 62*, 8, 1457–1506.

SCHMIDT-WEIGAND, F., KOHNERT, A., AND GLOWALLA, U. 2010. A closer look at split visual attention in system-and self-paced instruction in multimedia learning. *Learning and Instruction 20*, 100–110.