Ocular-based automatic summarization of documents: is re-reading informative about the importance of a sentence?

Orlando Ricciardi, Giovanni Serra, Federica De Falco, Piero Maggi, & Francesco Di Nocera
Sapienza University of Rome
Italy

Abstract

Automatic document summarization (ADS) has been introduced as a viable solution for reducing the time and the effort needed to read the ever-increasing textual content that is disseminated. However, a successful universal ADS algorithm has not yet been developed. Also, despite progress in the field, many ADS techniques do not take into account the needs of different readers, providing a summary without internal consistency and the consequent need to re-read the original document. The present study was aimed at investigating the usefulness of using eye tracking for increasing the quality of ADS. The general idea was of that of finding ocular behavioural indicators that could be easily implemented in ADS algorithms. For instance, the time spent in re-reading a sentence might reflect the relative importance of that sentence, thus providing a hint for the selection of text contributing to the summary. We have tested this hypothesis by comparing metrics based on the analysis of eye movements of 30 readers with the highlights they made afterward. Results showed that the time spent reading a sentence was not significantly related to its subjective value, thus frustrating our attempt. Results also showed that the length of a sentence is an unavoidable confounding because longer sentences have both the highest probability of containing units of text judged as important, and receive more fixations and re-fixations.

Introduction

Summarization is a strategy used to understand and store knowledge (Anderson & Armbruster, 2000). The goal of a summary is to produce a document shorter than the original by eliminating unnecessary information, allowing the readers to optimize their use of time and cognitive effort (Renkl & Atkinson, 2007) and to organize the text in a structure that facilitates comprehension (Leopold et al., 2013). The activity of rewriting text is the last phase in the process of summarization. Indeed, when an individual reads text to study it, s/he proceeds with a quick first reading, then determines the main contents and, finally, rewrites them into a new, shorter document (Flower & Hayes, 1980; Taylor & Beach, 1984; Wittrock & Alesandrini, 1990). An essential value of a summary is that it reflects precisely what the reader wants to learn about a topic. On the other hand, it is evident that the activity of providing summaries requires time and cognitive effort, especially if the text is

In D. de Waard, F. Di Nocera, D. Coelho, J. Edworthy, K. Brookhuis, F. Ferlazzo, T. Franke, and A. Toffetti (Eds.) (2018). Proceedings of the Human Factors and Ergonomics Society Europe Chapter 2017 Annual Conference. ISSN 2333-4959 (online). Available from http://hfes-europe.org

lengthy. Such problems are made worse by the increasing amount of electronic information available online and the consequent need to manage it quickly (for a review see Eppler & Mengis, 2004). ADS represents a partial solution to this problem by allowing the creation of summaries in a few seconds, by selecting the essential contents of a text (Gupta & Lehal, 2010).

Research in this field started with the interest in the production of abstracts for technical documentation (Saggion & Poibeau, 2013). Particularly, the first attempt to use ADS was in 1958, when Luhn proposed an algorithm that employs the frequency of a word to measure sentence relevance, leading to the first and most straightforward approach for creating a summary. Subsequently, more complex strategies have considered syntactic analysis of the text (Climenson et al., 1961), grammatical rules of discourse construction, and the semantic relationships among words and sentences (Mani & Maybury, 1999). In general, these approaches have used statistical techniques to extract one or more phrases to provide a summary (Paice, 1990) and are usually divided into two categories: abstractive and extractive methods (Hahn & Mani, 2000). Abstractive methods use linguistic approaches to identify the central concepts and produce a shorter text that may include new sentences, not stated in the original version (Erkan & Radev, 2004). Alternatively, extractive methods focus on statistical analysis of the text's features such as the unit's location in the source text, how often it occurs, the appearance of cue phrases, and statistical significance metrics (Hahn & Mani, 2000). This class of techniques attributes a weight or a score to each different word and sentence and uses statistical analysis to integrate linguistic features (word/phrase frequency, location of cue words) into a shorter document (Kyoomarsi et al., 2008). The fundamental principle is that the most frequent or the better-positioned content in a sentence is considered the most important. Even if these approaches are easy to implement, several issues limit their efficiency. The main problem is that this approach avoids analysing the text`s meaning, providing a summary that may be incomplete or without internal coherence (Hahn & Mani, 2000). For example, if two successive sentences explain different aspects of the same concept, and if only one of them is extracted, a re-analysis of the text might be required. It is worth noting that many studies have investigated this topic, but rarely have the human factors of summarization behaviour been examined (Xu et al., 2009).

Human Factors in summary evaluation

Summary evaluations have been discussed since the late 1990s (Jones & Endres-Niggemeyer, 1995; Mani et al., 2002). Automatic summary evaluation methods can be divided into “intrinsic” and “extrinsic” (Jones & Galliers 1996; Hahn & Mani, 2000). Intrinsic evaluations methods are based on the characteristics of the summary and do not consider the final user. These assessments focus on the consistency between different parts of the text, on the correspondence between the weights assigned to original sentences and extracted sentences and on the information reported into the summary compared to that in the original version. Alternatively, in extrinsic evaluation methods, the final user is the centre of the evaluation process. This technique evaluates how much the summary responds to the user's needs, considering readability, relevance and efficiency of the review based on a query.

Both intrinsic and extrinsic techniques can be automated or manual. In automated approaches, the evaluation consists of a comparison of the summary with one or more reference summaries (Saggion & Poibeau, 2013). Automatic procedures are better suited to extractive methods, whereas evaluations of summaries generated with abstract methods can be made only with a manual approach due to the difficulty of interpreting the meanings of new sentences (Saggion & Poibeau, 2013). In manual procedures, a team of different users evaluates a summary considering various features, such as style, grammar, content, readability, etc. These types of evaluations are often required as benchmarks, even in automated evaluation methods, but they are vulnerable to user subjectivity.

In summary, all ADS methods include the measurement of two fundamental properties: the Compression Ratio and the Retention Ratio. The first refers to the length of the summary relative to that of the original text, whereas the retention ratio indicates how much information from the original version has been retained in the summary (Mani & Maybury, 1999). Even though a significant amount of literature has been provided on this topic, the problem is still far from being solved. Thus, it is possible that a human-cantered perspective could help in addressing the issue. Our idea is to try to improve the quality of automatic summarization techniques by integrating a subjective behavioural indicator of importance into the extractive methods. Particularly, eye movements made during initial reading activity could reflect the reading strategy involved in detecting essential aspects of a text.

Eye behaviour in reading activity

The relationship between eye movements and attention has been widely studied, and several approaches have been proposed to describe it. The Premotor Theory (e.g., Rizzolatti et al., 1987; Rizzolatti, Riggio, & Sheliga, 1994) for example, suggests that attention and eye movements rely on the same brain structures. Furthermore, the Eye-Mind hypothesis (Just & Carpenter, 1984) advises a strong correlation between gazes and cognitive processes. Other studies have found that attention and saccades depend on the same mechanisms involved in spatial attention and in saccade orientation (Shepherd et al., 1986; Kowler et al., 1995; Kowler, 1996). In line with these results, Hoffman and Subramaniam (1995) found that subjects have difficulty in moving their eyes to one location and attending to another, even when instructed to do so and, in contrast, that making a saccade to an area improves the detectability of information presented in that location. In this framework, eye movements could be used to detect reading behaviour indicators that could be used as weights to select the information that will be included into the summary.

The availability of new eye-tracking technology allows us to gain a deep understanding of the eye movement behaviour during reading (Clifton et al., 2016; Radach & Kennedy, 2013; Rayner, 1975, 1978; Rayner & Pollatsek, 1987). Notably, several studies have reported the details of saccades, fixations, skipping and re-fixations (Liversedge et al., 2011; O'Regan & Ltvy-Schoen, 1987; O'Regan et al., 1984; Pynte, 1996; Pollatsek & Rayner, 1990; Reichle et al. 2003). For instance, there is a consensus on the variability of saccades (20-50 milliseconds) and fixation durations (200-500 milliseconds) due to the relation between oculomotor system behaviour and comprehension processing difficulties (Reichle et al. 2003). The most important factors affecting fixation durations are word length, frequency, age of acquisition, predictability (how predictable a word is from the context of a sentence) and similarity with other words (Ehrlich & Rayner, 1981; O'Regan, Levy-Schoen, Pynte & Brugaillère, 1984; Balota et al., 1985; Inhoff & Rayner, 1986; Kliegl et al., 2004; Hyönä , 2011; Rayner, 2009). Moreover, the duration of fixations may include the encoding of the antecedent word ("spillover effect"; Reichle et al., 2003), or the encoding of the successive word ("preview benefit"; Inhoff et al., 2000; Schroyens et al., 1999). About 10-15% of saccades are called "regressions" because the eyes move back to a part of a text that has been already inspected (see Rayner, 2009). These movements may be due to several factors such as: the correction of oculomotor errors (see Bicknell & Levy, 2011; O'Regan, 1990) for searching for the "optimal viewing position" (O'Regan, 1990; Brysbaert & Nazir, 2005), or difficulties in linguistic processing (Reichle et al., 2003).

Although a large body of research has been conducted with eye tracking during reading, few studies have tried to use eye-movement related metrics in ADS. Xu et al. (2009) for example, assume that the amount of time that a reader spends on a word is related to its importance in the comprehension process of the entire text. Following this reasoning, they inserted the "duration of fixations" into automatic summarization software as a criterion to determine which sentences to include in the summary. Although their results showed some superiority over other automatic summarization software, some issues remain. One issue is the relationship between attention and fixation duration. Several studies have indicated that fixation time on a word does not necessarily reflect its importance or the depth of cognitive processing. Indeed, long fixation time might also reflect difficulty in processing both the word fixated and the information derived from words in parafoveal vision (Kennedy & Pynte, 2005; Kliegl et al., 2006). The lack of attention paid by Xu and co-workers to the variety cognitive processes potentially affecting fixation times has been criticized by Buscher et al. (2012) in a more recent survey. In their study, the authors investigated the relationships among the following variables: "coherently read text length", considered as the length of text in characters that has been read consistently without skipping any text; the "thorough reading ratio", computed as the amount of text that has been detected as having been read divided by the amount of reading or skimmed text; the "regression ratio", i.e. the ratio between the number of regressions made on a single Area of Interest (AOI) divided by the total number of saccades received from that AOI, and the "mean forward saccade length", calculated as the average length of progressive saccades. The authors found that "coherently read text length," "thorough reading ratio" and "regression ratio" increased with perceived relevance of the text, but "mean forward saccade length" decreased with perceived relevance. These results suggest that essential sentences are the recipients of more accurate reading. Also, more important paragraphs and phrases are more frequently inspected by the reader. Finally, Buscher et al. (2012) found that fixation duration, as predicted in the literature, was not related to the importance of a sentence or paragraph. Despite the scientific contribution of this research, the results are not yet conclusive and satisfactory to produce more consistent automatic summarization algorithms.

Study

In the present study, we used eye tracking with the aim to improve the quality of ADS techniques. Thanks to the evolution in eye tracking technology and data analysis methods, we aimed to collect information during a reading task to be used to provide an index of importance attributed by the reader to a sentence. This measure could then be used to improve the quality of automatic summaries by tailoring the summary to the reader's goals. Specifically, we suggested that the study of eye movements during a natural reading activity could allow identification of the reading strategy used to create a summary. As indicators, we have considered first fixation duration as the time spent on the first reading of a sentence and re-fixation duration as the time spent on the second reading of a sentence (explained in more detail in the method section). We used the highlights made on a printed version of the same text as a measure of the subjectively perceived importance (Nist & Hogrebe, 1987; Peterson, 1991). The research hypothesis was that the time spent in re-reading a sentence reflects the subjectively perceived importance (SPI) of that sentence. Along with this assumption, we expected that a sentence receiving longer re-fixations also should receive more highlights. To test this hypothesis, eye fixation behaviour recorded during a screen reading task was compared with the importance attributed to specific sentences by observers who underlined parts of its printed version (Nist & Hogrebe, 1987; Peterson, 1991).

Participants

Thirty university students (25 females, mean age = 26.4; sd = 4.5) volunteered to participate. All had normal or corrected-to-normal vision.

Materials and Method

We used a magazine article for the study, with the aim to involve the participants in the reading activity. The text chosen is the official Italian translation of the article "Academy Fight Song" by Thomas Frank (available at the web address https://thebaffler.com/salvos/academy-fight-song). The text was divided into 41 pages and presented as slideshow. Each sentence was attributed to an AOI for collecting the fixations with the Tobii Studio software, allowing counting the fixations for each AOI. For example, Figure 1 shows the editing of different AOIs in the Tobii Studio software *(a)* and the version read by the participants during the experimental session *(b)*.

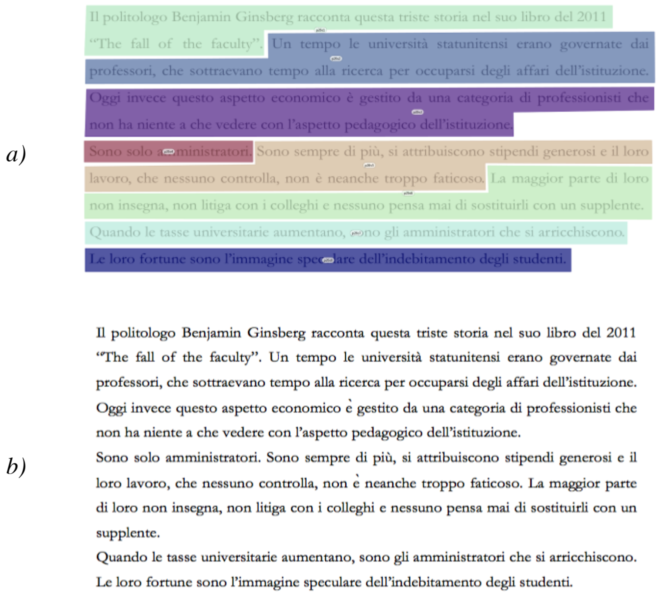

Figure 1. a) A specific Area of Interest was assigned to each sentence in the text; b) The version of the text read by the participants during the experimental session.

First fixations (FF) were defined as the early exploration on the "n" AOI until the eyes moved on to the "n + 1" AOI. Then, each backward movement on the "n" AOI was considered a re-fixation (re-reading, RR; Figure 2). The total number and durations of FFs and RRs were weighted according to the length of the AOI, to avoid biasing the data by the number of characters present in the sentences. The X2-30 eye tracker system (Tobii, Sweden) was used to record eye movements during the reading activity.

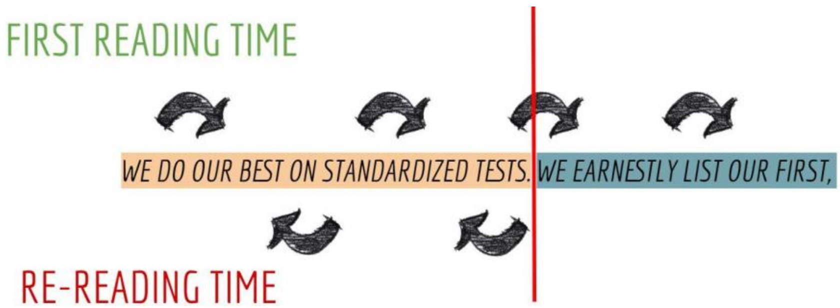

Figure 2. First reading was considered until a backward movement happened from the "n + 1" AOI to the "n" AOI. All successive fixations on the "n" AOI were considered as re-reading.

Procedure

The experiment consisted of two phases: in the first, participants were asked to read a magazine article on a 17" screen while they were positioned at about 60 cm from the display, and the text was displayed in a full-screen mode, to facilitate a comfortable reading. Before the experiment, a 9-point calibration was performed. Subjects were instructed in using the spacebar for moving to the next slide. The second phase of the experiment took place after a week, with the same subjects. The task consisted of reading the same version of the magazine article in a printed version. We asked them to highlight the most important concepts contained in the text with the objective of collecting an indicator of the subjective importance attributed to the sentences. In both phases, reading comprehension was assessed with a brief structured interview (i.e., "What problem is discussed in the article that you have just read?"; "Why are American students willing to apply for a loan to attend a college or a university?").

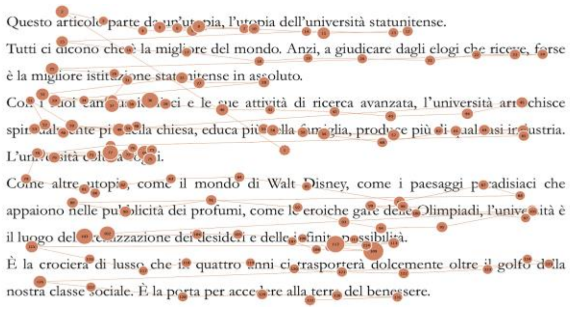

Figure 3. Scan path: saccades (segments) and fixations (spheres) recorded during the reading task.

Data analysis and results

The data used for the analyses were the numbers and durations of fixations and re-fixations directed to each sentence (AOI). The highlights collected on the printed version of the article were used as a subjective measure of perceived importance in the analyses. AOIs were classified into four categories ("very low," "low," "high," "very high"), according to the sum of highlights received from all the subjects (quartiles were considered for classifying the sentences).

Data analyses showed a high positive correlation between the number and the duration of fixation both when the subject read for the first time ($r = .99$) both in re-reading ($r = .97$). Due to this high correlation, we have decided to further analyse only duration. The correlation between the number of highlights for each sentence and the re-reading time was significant but low ($r = .21$). Also, the correlation with the first reading time was found to be significant ($r = .31$). Moreover, a significant positive correlation was found between the highlights and the sentence's length ($r = .33$).

Due to this correlation, this variable has been included as a covariate in an ANCOVA design, where first reading time and re-reading time were used as dependent variables and the number of highlights (category) as a factor. The analysis resulted in the absence of the primary effect of both the variables, showing no difference in reading or reading time depending on the highlights marked on each group of sentences [$F_{3,205} = .45$, $p > .05$]. Indeed, a positive relation was found for the covariate (Figure 4 and 5).

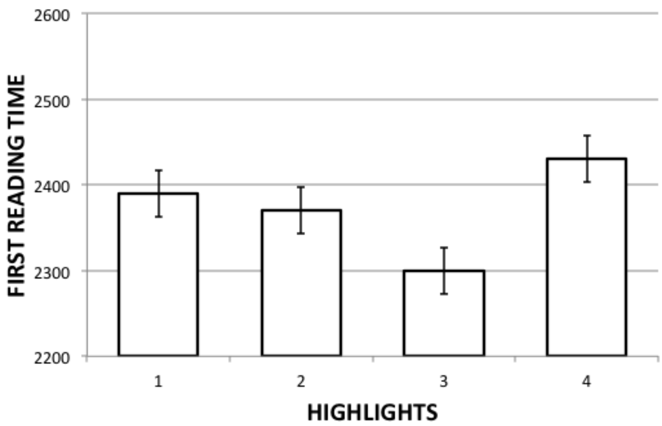

Figure 4. First Reading time (in milliseconds) by highlighting (quartiles indicate highlighting increment and therefore the importance of the sentence); bars represent the 95% confidence interval.

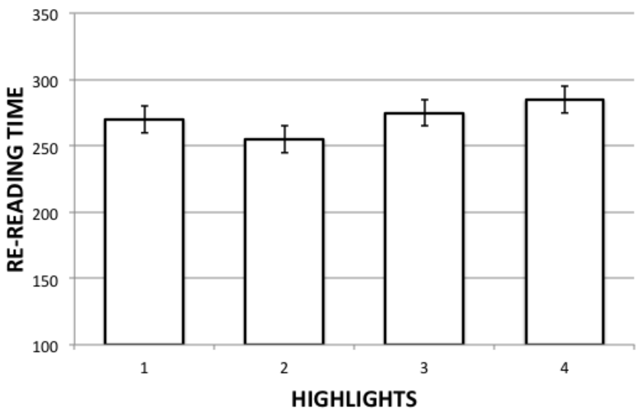

Figure 5. Re-Reading time (in milliseconds) by highlighting (quartiles indicate highlighting increment and therefore the importance of the sentence); bars represent the 95% confidence interval.

Therefore, the sentence's length was used as a dependent variable in a one-way ANOVA design, using the highlights as a factor. The analysis confirmed the significant relationship of this variable [$F_{3,206} = 10.61$, $p < .0001$; Figure 6).

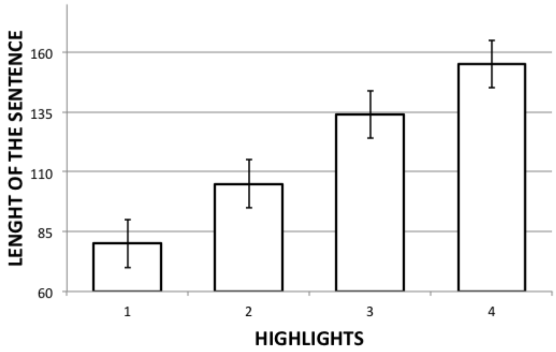

Figure 6. Length of the sentence (in characters) by highlighting (quartiles indicate highlighting increment and therefore the importance of the sentence); bars represent the 95% confidence interval.

Discussion

The objective of the present study was to improve ADS techniques by devising an indicator of the subjective importance of text based on the analysis of eye movement behaviour. Therefore, we compared eye movements and highlighting strategies to find a relation between a behavioural and an individual attribution of salience across the text. Indeed, each sentence has been considered as an AOI and the time spent on it was correlated with the number of highlights received. Highlights were considered an indicator of personal importance (Nist & Hogrebe, 1987; Peterson, 1991), as each reader was free to select the most critical concepts based on his or her previous knowledge and the reader`s objectives. At the same time, eye movement behaviour is deeply involved in cognitive aspects of reading activity (Liversedge et al., 2011; O'Regan & Ltvy-Schoen, 1987; O'Regan et al., 1984; Pollatsek & Rayner, 1990; Rayner et al., 1998; Reichle, 2003).

The idea of integrating a behavioural indicator into an extractive summarization technique is due to the current problems with these classes of methods, since they are usually unable to provide a summary that can satisfy the reader's goals (Hahn & Mani, 2000). Indeed, these algorithms are easy to implement because they select sentences with higher scores, depending on some indicators such as word frequency or word location (Kyoomarsi et al., 2008). However, they do not analyse text meaning and often provide a summary with poor internal coherence because of the loss of crucial concepts (Hahn & Mani, 2000). Eye movement data could offer a subjective weight that can be used to tailor the summary according to the reader's goal, advancing the utility of the extractive methods of summarization. If this relation were confirmed, the applicative rate would be enormous.

Summarizing a text is a common strategy to study and store information (Anderson & Armbruster, 2000), and automatic summarization is used to reduce time and cognitive efforts needed to produce a summary (Renkl & Atkinson, 2007). This is even more important considering the enormous amount of information available online and the need to manage it quickly (Eppler & Mengis, 2004).

The primary hypothesis of our study was that the duration of the regressions made during reading reflects the subjectively perceived importance to the reader. More specifically, we hypothesized that higher re-reading times should correspond with more highlighted sentences, whereas lower re-reading times should be associated with sentences that received fewer highlights.

The results obtained did not confirm our initial hypothesis. The correlation matrix showed a very high correlation between time and number of fixation (both in FR and in RR), so we decided to use the average time of fixation as dependent variable for the analyses. Only a weak correlation was found between RR and highlights. At the same time, the relation between FR measures and the highlights was also low. However, a significant correlation was found between the length of sentences and the number of highlights. It is worth noting that the number of characters in each phrase was balanced, as the FR and RR on each AOI were divided by the number of characters contained in that AOI. FR and RR were used as dependent variables in an ANCOVA design, using the proportion of highlights (Very Low, Low, High and Very High categories) as a factor and the sentence`s length as a covariate. The main effects were nonsignificant, and no interaction effect was found, suggesting that the time spent on reading or rereading the text was not related to the highlighting strategy. Indeed, a significant effect of the covariate was found in both analyses, confirming the relationship between the sentence's length and the eye movement behaviour.

Although the results of this study are far from being conclusive, they seem to disconfirm the hypothesis that the time spent on a sentence reflects its relative importance. This effect was found by Buscher et al. (2012), for example. In their study, the authors noted that essential contents induce more precise eye movements and a higher probability of re-fixation, while contents perceived as not relevant are more related to "skimming" behaviour; that is, a higher likelihood of scanning the text very quickly, skipping many words and without re-fixation. The difference between our results and those from Buscher and colleagues (2012) could be due to several factors as, for instance, the text's language. Italian and English writing styles have different linguistic structures, and one of the main aspects is the sentence`s length, usually shorter in the English language. Moreover, in our study participants did not have to provide a summary of the text, but only to perform a verbal assessment of text comprehension. It is possible that the goal of the task determined a lower level of effort in the second task, and that the subjects gave more importance only to the more extended sentences.

Limitation and further research

The study has some limitations, including the assumption that eye movement regressions and re-reading time reflects the importance given to a sentence. Indeed, re-reading should also reveal the reader's difficulties in understanding a sentence (Rayner et al., 2006). This hypothesis was not explored in the present study but, considering our results, needs to be addressed in further research on this topic.

An issue lies in the instrumentation used to detect eye movements; i.e., a low-cost eye-tracker (sampling rate of 30Hz) that might have led to some errors in measuring the eye movement behaviour during reading. It is possible that some fixations and saccades were not detected, or they had been assigned to an incorrect AOI. Of course, a higher sampling rate would make easier to collect these kinds of data (Holmqvist et al., 2011), but we observed that our effort for improving ADS techniques by making use of eye-tracking measures is one of the first studies of its kind in this area. Another limitation could be attributed to the experimental design, in which several factors might have altered ocular behaviour. The absence of pauses during the task can be considered a problem, as the subjects reported being tired at the end of the recording session (that lasted from 30 to 40 minutes). Indeed, we had chosen not to allow pauses during the reading activity to avoid the need to re-calibrate the eye tracker. That limitation might have led to a loss in ecological validity. Regarding the data analysis, even though we had tried to weigh and normalize the collected measures, some features that may change the subjective reading behaviour have not been considered. We did not control the word frequency and the word predictability, for example. We were aware that these variables are considered significant in studies that focus on eye movements and reading. However, we should emphasize that our intent was not to study reading activity itself but to analyse the relationship between ocular behaviour during reading and a subsequent highlighting strategy during the study of a document. For this objective, we thought that the best way to investigate this relationship was to use a natural text, accepting loss of control of its structure. This choice was considered trivial to maintain the ecological validity of the study and to enable application of our results to future research and developing ADS technology.

Overall, additional research is required to better understand the relationship between re-reading times and perceived importance of text segments. Future experimental designs should include some modifications to reduce the impact of the described limitations. For example, splitting the reading task into shorter sessions and presenting the subjects with different types of documents, could help to control effects due to fatigue and document type. Additionally, all the features that influence reading times should be limited or controlled.

References

Anderson, T.H., & Armbruster, B.B. (2000). Studying. In M.L. Kamil, P.B. Mosenthal, P.D. Pearson, and R. Barr (Eds.), *Handbook of reading research, Volume III* (pp. 657 - 679). Mahwah, NJ: Erlbaum.

Balota, D.A., Pollatsek, A., & Rayner, K. (1985). The interaction of contextual constraints and parafoveal visual information in reading. *Cognitive Psychology, 17,* 364-390.

Bicknell, K., & Levy, R. (2011). Why readers regress to previous words: A statistical analysis. In L. Carlson, C. Holscher, and T. Shipley (Eds.), *Proceedings of the 33rd Annual Meeting of the Cognitive Science Society* (pp. 931–936). Austin, TX: Cognitive Science Society.

Brysbaert, M., & Nazir, T. (2005). Visual constraints in written word recognition: evidence from the optimal viewing position effect. *Journal of Research in Reading, 28,* 216-228.

Buscher, G., Dengel, A., Biedert, R. & Elst. L.V. (2012). Attentive documents: Eye tracking as implicit feedback for information retrieval and beyond. *ACM Transactions on Interactive Intelligent Systems, 1*, 9.

Clifton, C., Ferreira, F., Henderson, J.M., Inhoff, A.W., Liversedge, S.P., Reichle, E.D., & Schotter, E.R. (2016). Eye movements in reading and information processing: Keith Rayner's 40year legacy. *Journal of Memory and Language, 86,* 1-19.

Climenson, W.D., Hardwick, N.H., & Jacobson, S.N. (1961). Automatic syntax analysis in machine indexing and abstracting. *American Documentation*, 178-183.

Ehrlich, S.E, & Rayner, K. (1981). Contextual effects on word perception and eye movements during reading. *Journal of Verbal Learning and Verbal Behavior, 20,* 641-655.

Eppler, M.J., & Mengis, J. (2004). The concept of information overload: A review of literature from organization science, accounting, marketing, MIS, and related disciplines. *The information society, 20,* 325-344.

Erkan, G., & Radev, D. R. (2004). Lexrank: Graph-based lexical centrality as salience in text summarization. *Journal of Artificial Intelligence Research, 22,* 457-479.

Flower, L., & Hayes, J. (1980). The dynamics of composing: making plans and juggling constraints. In L. Gregg, and E. Steinberg (Eds.), *Cognitive processes in writing* (pp. 31-50). Hillsdale, NJ: Lawrence Erlbaum Associates.

Gupta, V., & Lehal, G.S. (2010). A survey of text summarization extractive techniques. *Journal of emerging technologies in web intelligence, 2*, 258-268.

Hahn, U., & Mani, I. (2000). The challenges of automatic summarization. *Computer, 33*, 29-36.

Hoffman, J.E., & Subramaniam, B. (1995). The role of visual attention in saccadic eye movements. *Attention, Perception, & Psychophysics, 57*, 787-795.

Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & Van de Weijer, J. (2011). *Eye tracking: A comprehensive guide to methods and measures*. OUP Oxford.

Hyönä, J. (2011). Foveal and parafoveal processing during reading. In S.P. Liversedge, I.D. Gilchrist, and S. Everling (Eds.), *The Oxford Handbook of Eye Movements* (pp. 819 – 837). Oxford, England: Oxford University Press.

Inhoff, A.W., & Rayner, K. (1986). Parafoveal word processing during eye fixations in reading: Effects of word frequency. *Perception & Psychophysics, 40,* 431–439.

Inhoff, A.W., Starr, M., & Shindler, K. (2000). Is the processing of consecutive words during eye fixations in reading strictly serial? *Perception & Psychophysics, 62,* 1474-1484.

Jones, K.S., & Endres-Niggemeyer, B. (1995). Automatic summarizing. *Information Processing & Management, 31*, 625-630.

Jones, K.S., & Galliers, J. (1996). Evaluating natural language processing systems: An analysis and review. *Natural Language Engineering, 4,* 175-190.

Just, M.A., Carpenter, P.A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review, 87*, 329-354.

Kennedy, A., & Pynte, J. (2005). Parafoveal-on-foveal effects in normal reading. *Vision Research, 45,* 153–168.

Kyoomarsi, F., Khosravi, H., Eslami, E., Dehkordy, P.K., & Tajoddin, A. (2008). Optimizing text summarization based on fuzzy logic. In *Computer and Information Science, 2008. ICIS 08. Seventh IEEE/ACIS International Conference on* (pp. 347-352). IEEE.

Kliegl, R., Grabner, E., Rolfs, M., & Engbert, R. (2004). Length, frequency, and predictability effects of words on eye movements in reading. *European Journal of Cognitive Psychology, 16*(1-2), 262-284.

Kliegl, R., Nuthmann, A., & Engbert, R. (2006). Tracking the mind during reading: the influence of past, present, and future words on fixation durations. *Journal of experimental psychology: General, 135*, 12.

Kowler, E. (1996). Cogito ergo moveo: cognitive control of eye movement. In *Exploratory vision* (pp. 51 - 77). Springer New York.

Kowler, E., Anderson, E., Dosher, B., & Blaser, E. (1995). The role of attention in the programming of saccades. *Vision research, 35*, 1897-1916.

Khosravi, H., Eslami, E., Kyoomarsi, F., & Dehkordy, P. (2008). Optimizing text summarization based on fuzzy logic. *Computer and Information Science,* 121-130.

Leopold, C., Sumfleth, E., & Leutner, D. (2013). Learning with summaries: Effects of representation mode and type of learning activity on comprehension and transfer. *Learning and Instruction, 27,* 40-49.

Liversedge, S.P., Gilchrist. I.D. & Everling, S. (2011). *The Hoxford Handbook Of Eye Movements*. New York: Oxford University Press Inc.

Mani, I., & Maybury, M. T. (1999). *Advances in automatic text summarization, Volume 293*. Cambridge, MA: MIT press.

Mani, I., Klein, G., House, D., Hirschman, L., Firmin, T., & Sundheim, B. (2002). SUMMAC: a text summarization evaluation. *Natural Language Engineering, 8*, 43-68.

Nist, S.L., & Hogrebe, M.C. (1987). The role of underlining and annotating in remembering textual information. *Literacy Research and Instruction, 27*, 12-25.

O'Regan, J.K. (1990). Eye movements and reading. *Reviews of oculomotor research, 4,* 395.

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

O'Regan, J.K., & Ltvy-Schoen, A. (1987). Eye movement strategy and tactics in word recognition and reading. In M. Coltheart (Ed.), *Attention and Performance: Vol.12. The psychology of reading* (pp. 363-383). Hillsdale, NJ: Erlbaum.

Paice, C.D. (1990). Constructing literature abstracts by computer: techniques and prospects. *Information Processing & Management, 26*, 171-186.

Peterson, S.E. (1991). The cognitive functions of underlining as a study technique. *Literacy Research and Instruction, 31*, 49-56.

Pollatsek, A., & Rayner, K. (1990). Eye movements and lexical access in reading. In D.A. Balota, G.B. Flores d'Arcais, and K. Rayner (Eds.), *Comprehension processes in reading* (pp. 143 - 164). Hillsdale, NJ: Erlbaum.

Pynte, J. (1996). Lexical control of within-word eye movements. *Journal of Experimental Psychology: Human Perception and Performance, 22,* 958-969.

Radach, R., & Kennedy, A. (2013). Eye movements in reading: Some theoretical context. *The Quarterly Journal of Experimental Psychology, 66*, 429-452.

Rayner, K. (1975). Parafoveal identification during a fixation in reading. *Acta Psychologica, 39*, 271-281.

Rayner, K. (1978). Eye movements in reading and information processing. *Psychological Bulletin, 85,* 618-660.

Rayner, K. (2009). The Thirty Fifth Sir Frederick Bartlett Lecture: Eye movements and attention in reading, scene perception, and visual search. *The Quarterly Journal of Experimental Psychology, 62,* 1457–1506.

Rayner, K. & Pollatsek, A. (1987) Eye movements in reading: A tutorial review. *Attention and performance, 12I,* 327–362.

Rayner, K., Chace, K.H., Slattery, T.J., & Ashby, J. (2006). Eye movements as reflections of comprehension processes in reading. *Scientific studies of reading, 10*, 241-255.

Rayner, K., Reichle, E.D., & Pollatsek, A. (1998). Eye movement control in reading: An overview and model. *Eye guidance in reading and scene perception,* 243-268.

Reichle, E.D., Rayner, K., Pollatsek, A. (2003). The E-Z Reader model of eye movement control in reading Comparison to other models. *Brain and Behavioral Sciences, 26,* 445–476.

Renkl, A., & Atkinson, R.K. (2007). Interactive learning environments: Contemporary issues and trends. An introduction to the special issue. *Educational Psychology Review, 19,* 235-238.

Rizzolatti, G., Riggio, L., & Sheliga, B.M. (1994). Space and selective attention. *Attention and performance XV, 15,* 231-265.

Rizzolatti, G., Riggio, L., Dascola, I., & Umiltá, C. (1987). Reorienting attention across the horizontal and vertical meridians: evidence in favor of a premotor theory of attention. *Neuropsychologia, 25*, 31-40.

Saggion, H., & Poibeau, T. (2013). Automatic text summarization: Past, present and future. In *Multisource, multilingual information extraction and summarization* (pp. 3 - 21). Springer Berlin Heidelberg.

Schroyens, W., Vitu, F., Brysbaert, M., & d'Ydewalle, G. (1999). Eye movement control during reading: Foveal load and parafoveal processing. *Quarterly Journal of Experimental Psychology, 52*, 1021-1046.

Shepherd, M., Findlay, J.M., & Hockey, R.J. (1986). The relationship between eye movements and spatial attention. *The Quarterly Journal of Experimental Psychology, 38*, 475-491.

Taylor, B.M., & Beach, R.W. (1984). The effects of text structure instruction on middle-grade students comprehension and production of expository text. *Reading Research Quarterly, 19,* 134-146.

Wittrock, M.C., & Alesandrini, K. (1990). Generation of summaries and analogies and analytic and holistic abilities. *American Educational Research Journal, 27,* 489-502.

Xu, S., Jiang, H. & Lau, F.C. (2009). User-oriented document summarization through vision-based eyetracking. In *Proceedings of the 13th International Conference on Intelligent User Interfaces,* (pp. 7-16). New York: ACM.