Déjà Vu: Eye Movements in Repeated Reading

Yoav Meiri[1] **(meiri.yoav@campus.technion.ac.il)**
Yevgeni Berzak[1,2]**(berzak@technion.ac.il)**
[1] Faculty of Data and Decision Sciences, Technion - Israel Institute of Technology, Haifa, Israel
[2] Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, USA

Abstract

From cooking recipes to novels and scientific papers, we often read the same text more than once. How do our eye movements in repeated reading differ from first reading? In this work, we examine this question at scale with L1 English readers via standard eye-movement measures and their sensitivity to linguistic word properties. We analyze consecutive and non-consecutive repeated reading, in ordinary and information-seeking reading regimes. We find sharp and robust reading facilitation effects in repeated reading, and characterize their modulation by the reading regime, the presence of intervening textual material, and the relevance of the information to the task across the two readings. Finally, we examine individual differences in repeated reading effects and find that their magnitude interacts with reading speed, but not with reading proficiency. Our work extends prior findings, providing a detailed empirical picture of repeated reading which could inform future models of eye movements in reading.[1]

Keywords: eye movements in reading, repeated reading

Introduction

Analysis of eye movements in reading is a key methodology in the study of human language processing. Research that uses this methodology relies primarily on experimental paradigms that presuppose that the reader has not previously encountered the experimental text (Rayner, 1998; Kliegl et al., 2004). This assumption is required for many research questions. However, our daily experience of language comprehension includes not only novel, but also previously processed linguistic stimuli. In the domain of reading, multiple readings of the same text, which can occur at different repetition intervals, are a common scenario. For example, in education, 84% of US college students listed repeated reading as a learning strategy they use (Karpicke et al., 2009). Despite its ubiquity, repeated reading received limited attention in the psycholinguistic literature.

In this work, we argue that repeated reading is an important building block in the development of more comprehensive empirical accounts and cognitive theories of language processing during reading. We take several steps to advance research on eye movements in repeated reading which examine this phenomenon under different reading regimes, amounts of intervening text between readings, divisions of task relevant and task irrelevant information across readings, and from the perspective of individual differences. Our analyses are enabled by OneStop (Malmaud et al., 2020), a broad coverage English dataset of eye movements in reading, with 360 adult English L1 participants, 2,110,632 word tokens with eye-tracking data in first reading, and 422,167 in repeated reading. This dataset is the largest resource on repeated reading to date, both in the number of participants and the amount of eye-tracking data.

Using this dataset, we analyze an array of global eye movement measures in repeated reading. We further examine the responsiveness of participants' reading behavior to predictability, frequency and word length, three word properties that were shown to be robust predictors of reading times (Kliegl et al., 2004, among others). Departing from prior work, our analyses are conducted at different repetition intervals, ranging from consecutive article presentation to 10 articles between the readings, allowing to systematically examine the durability of the encoding of linguistic material throughout the experiment. We further compare these effects in an ordinary reading regime against the less commonly studied information-seeking regime (Hahn & Keller, 2023; Shubi & Berzak, 2023). Finally, we examine two key factors in individual differences in repeated reading, reading speed and reading proficiency.

Our analyses yield several results on repeated reading:

- In line with prior literature, we observe large and robust *reading facilitation*, exhibited in shorter reading times, and fewer fixations and regressions. Differently from Raney & Rayner (1995), we also find smaller word property effects.
- Facilitation effects tend to be larger in information seeking compared to ordinary reading.
- In ordinary reading, but not information seeking, repeated reading facilitation is moderately reduced in the presence of intervening material across readings.
- In information seeking, facilitation is modulated by the relevance of the information to the task, and by the similarity of the tasks across the two readings.
- Individual differences: facilitation effects are larger for slower readers, but do not depend on reading proficiency, as measured by reading comprehension performance.

Taken together, these results suggest that prior exposure to the text leads to reading that is highly resource efficient and simultaneously constrained by memory and the nature of the interactions with the text across readings.

[1] Code is available here.

Related Work

Prior work on eye movements in repeated reading focused primarily on benchmarking global eye movement measures. The main broad finding in this line of work is that such measures reflect reading facilitation. Hyönä & Niemi (1990) examine two repeated readings of the same text, an immediately consecutive reading, and a reading a week later. Across the two repeated readings they observe progressively shorter sentence reading times, fewer fixations per sentence, shorter average fixation duration and fewer regressions per sentence as compared to the first presentation of the text. In Raney & Rayner (1995) participants read two passages, each twice in succession. They find that the second reading is marked by shorter and fewer fixations, and longer saccades. Our analysis results for global eye movement measures are qualitatively consistent with the results obtained in these studies.

A key contribution of the current work is the analysis of word property effects on reading times in repeated reading. To our knowledge, the only study that previously examined this question with adult readers is Raney & Rayner (1995) mentioned above. This study included a target word frequency manipulation: high frequency versus low frequency. It found *no interaction* between the reading condition (first versus repeated) and the frequency effect size. Our work extends the frequency effect analysis of Raney & Rayner (1995) from binned values of target words to continuous values with broad corpus coverage, and further includes analyses of the effects of word predictability and length, which were shown to be key factors in determining reading times (Rayner et al., 2004; Kliegl et al., 2004; Rayner et al., 2011). Differently from Raney & Rayner (1995) we find a repeated reading–frequency interaction, and similar interactions for predictability and length.

Repeated reading has also been examined with respect to high-level semantic properties of the text, as well as the reader's goals. Hyönä (1995) use a similar experimental setup to Hyönä & Niemi (1990), to further demonstrate progressive reduction in the sensitivity of eye movements to the introduction of new topics in a second and third reading of a text. Closest to our study is Kaakinen & Hyönä (2007) who observe repetition benefits of similar magnitude for both perspective relevant and perspective irrelevant information in a second reading. Our analyses of information seeking are conceptually similar in the separation of task relevant and task irrelevant information, where we do find stronger facilitation effects for task irrelevant information.

We note that in education, repeated reading is used as a pedagogical technique during the acquisition of reading skills in children. A substantial body of work supports its effectiveness in improving reading fluency in children (Meyer & Felton, 1999; Faulkner & Levy, 1999; Teigen et al., 2001; Kuhn, 2004; Ardoin et al., 2008, among others). In this context, Foster et al. (2013) conducted an eye-tracking study with second grade children reading the same text 4 times consecutively. With each successive reading, they observe decreasing reading times for first fixation, gaze duration and total fixation, and fewer fixations and regressions. Differently from the results of Raney & Rayner (1995) with adult readers, frequency effects in children diminish over readings. Zawoyski et al. (2015) show similar results, modulated by the child's reading proficiency (high vs. low). Similarly to these results with children, our study with adults also shows interactions of word property effects with repeated reading. However, we find that in adults these effects are not modulated by reading proficiency, but rather by reading speed. Our study does not speak directly to the cognitive processes involved in repeated reading in the acquisition of reading skills in children, but could serve as an adult benchmark for comparison in such studies in the future.

Data and Experimental Setup

We use OneStop (Malmaud et al., 2020), a broad coverage English corpus of eye movements in reading, collected using an EyeLink 1000 Plus eye tracker (SR Research) at a sampling rate of 1000Hz. The dataset includes 360 L1 English participants reading materials from the OneStopQA corpus (Berzak et al., 2020), which includes a total of 30 Guardian articles with 162 paragraphs. The articles come in two difficulty level versions, an Advanced original Guardian version (19,444 words) and an Elementary version simplified for language learners (15,744 words). Each paragraph has 3 multiple choice questions. The textual span that is essential for answering a given question correctly, called the *critical span*, is manually annotated in the paragraph. Two of the three questions for each paragraph have the same critical span, and the remaining question has a distinct critical span.

In a single trial, participants read one paragraph in one of its two difficulty versions and then proceed to answer one question on a new screen, without the ability to return to the paragraph. The articles are divided into 3 experimental batches, with 10 articles (54 paragraphs) in each batch. Each participant reads a single 10-article batch in a random presentation order of the articles. Then, 2 articles are presented for a second time, with identical paragraphs and a different comprehension question for each paragraph. The experiment has two between-subjects reading regimes, ordinary reading (Gathering) and information seeking (Hunting) with an equal number of 180 participants in each regime. In the Hunting regime, participants are presented with the question (but not the answers) prior to reading the paragraph. In Gathering, participants see the question only after having read the paragraph. The mean experiment duration, excluding calibrations and breaks and including question previews (for Hunting), is 57 minutes in Gathering and 54 minutes in Hunting.

The article in position 11 is a second presentation of the article in position 10. The article in position 12 is a second presentation of one of the articles in positions 1 through 9. Thus, half of the repeated reading data captures consecutive repeated reading at the level of a full article, and the other half is repeated reading with intervening reading material, ranging from 2 to 10 articles. The data for each 60 participants in a combination of a 10-article batch and reading regime is counterbalanced such that each article appears 6 times in each of the positions 11 and 12. Further, the first presentation of articles repeated in position 12 is distributed across positions such that there is a total of 6 article appearances in each of the positions 2-9, with each article appearing at most once in each position, and a total of 12 article appearances in position 1. This experimental design corresponds to a total of 720 repeated presentations of articles with 3,888 paragraph trials and a total of 422,167 word tokens over which eye movements were collected, split equally between positions 11 and 12 and the two reading regimes. In first reading, there is a total of 3,600 article presentations, 19,438 paragraph trials and 2,110,632 word tokens with collected eye-tracking data.

Reading comprehension performance The mean reading comprehension accuracy in Gathering is 81.2% in first reading and 84.1% in repeated reading, which is higher ($p < 10^{-3}$)[2]. In Hunting, it is 86.9% in the first reading and 90.5% in repeated reading, which is also higher ($p < 10^{-6}$). In information seeking repeated reading, the comprehension accuracy is 91.2% when the critical spans of the questions in the two readings are identical, and 90.4% when the critical spans differ. We further note that participants' reading comprehension performance does not correlate with reading speed, as shown in Figure 1 in the Supplementary Material (SM, available here).

Global Eye Movement Measures

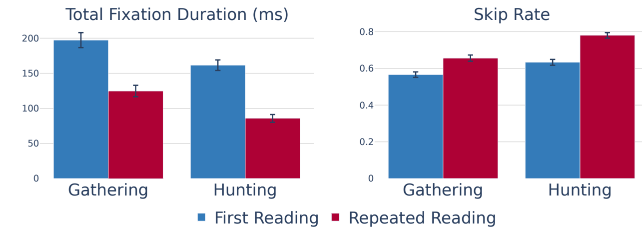
We first compare eye movements in first and second reading by examining the following eye movement measures:

- Total Fixation Duration (TF): the sum of all the fixation durations on a word. Words that were not fixated are assigned the value 0.

- First Pass Skip Rate (SR): the fraction of words that were skipped during the first pass.

The SM further includes analyses for First Fixation, first pass Gaze Duration, Gaze Duration, number of fixations per word, global Skip Rate ($TF = 0$) and Regression Rate.

Figure 1a presents the means of TF and SR for first and repeated text reading in the Gathering and Hunting reading regimes[3]. Compared to the first reading, in repeated reading we observe in each of the regimes substantially shorter TF and higher SR ($p < 10^{-58}$ for both measures and regimes)[4]. For SR, we further find a larger repeated reading effect in Hunting compared to Gathering ($p < 10^{-13}$)[5]. Figure 2a in the SM similarly shows facilitation effects in Gathering and in Hunting across all additional measures, with larger effects



(a)

(b)

Figure 1: Mean per word Total Fixation Duration and First Pass Skip Rate with 95% confidence intervals across ordinary reading (Gathering) and information seeking (Hunting) in first and repeated reading. (a) Overall (b) As a function of the article presentation position in the experiment. Positions 1–10: first reading. Position 11: a consecutive repeated reading of the article in position 10. Position 12: a non-consecutive repeated reading of one of the articles 1–9.

in Hunting in all measures except for Regression Rate and first pass Gaze Duration. These results are generally in line with prior findings on ordinary reading facilitation effects in repeated reading, and further demonstrate that such effects are also present and tend to be even larger in information seeking.

The experimental setup of OneStop enables us to go beyond the coarse distinction of first versus second reading examined in prior work, and address more fine-grained aspects of repeated reading. The first such aspect is the amount of intervening material between readings. Figure 1b breaks down the results in Figure 1a as a function of article position in the experiment. During the first reading, in both reading regimes, TF decreases ($p < 10^{-12}$ for both regimes). SR increases in Hunting ($p < 10^{-17}$) but not in Gathering ($p = 0.13$) as a function of the article position. Contrary to this trend, in Gathering TF *increases* in position 12 relative to position 11 ($p < 0.04$). A qualitatively similar trend reversal is observed for SR, albeit not significantly ($p = 0.38$). In Figure 2b in the SM, we see that the trend reversal from position 11 to 12 is highly consistent in Gathering across all the examined measures ($p < 0.045$ for all other measures). This suggests a reduction in reading facilitation when there is intervening material between the readings as compared to consecutive repeated reading.

Interestingly, this reversal is not observed in Hunting, where TF is shorter ($p < 0.02$) and SR is higher ($p < 10^{-3}$) in position 12 compared to position 11. Other measures in position 12 similarly follow the pattern in positions 1 through 11 (SM Figure 2b, $p < 0.02$ for all other measures except for first pass

[2] $correct \sim repeated + (repeated|subj) + (repeated|parag)$ applied to each reading regime (Hunting, Gathering).

[3] $RT \sim 1 + (1|subj) + (1|parag)$ applied to each combination of text presentation mode (first, repeated) and reading regime.

[4] $RT \sim repeated + (repeated|subj) + (repeated|parag)$ applied separately to each reading regime.

[5] Interaction term $repeated : regime$ in $RT \sim repeated * regime + (repeated|subj) + (repeated * regime|parag)$.

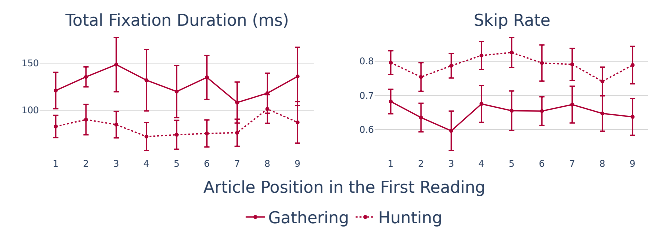


Figure 2: Mean per word Total Fixation Duration and First Pass Skip Rate with 95% confidence intervals *of the 12th article*, a repeated reading of one of the articles in positions 1-9, as a function of the article position in the first reading.

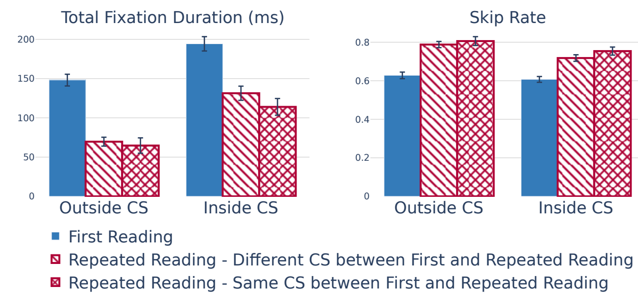


Figure 3: Mean per word Total Fixation Duration and First Pass Skip Rate with 95% confidence intervals *in information seeking (Hunting)* in first and repeated reading, within and outside the critical span (CS).

Gaze Duration where $p = 0.89$). These results suggest that differently from ordinary reading, the intervening material manipulation does not hinder reading facilitation in information seeking. This difference contributes to the overall larger reading facilitation in information seeking.

Can the precise article position during the first reading be traced down to global eye movement measures during the second reading? The answer to this question appears to be no. Figure 2 depicts repeated reading TF and SR of article 12 as a function of its first presentation position. A smaller article position corresponds to more intervening textual material and a larger time difference between the two readings. In both Gathering and Hunting, we observe no differences across positions. Figure 3 in the SM shows a similar result for the other measures. Thus, while there is a modest effect that distinguishes article-level consecutive and non-consecutive repeated reading in ordinary reading, simple global measures during repeated reading are not indicative of the precise number of articles that appeared between the two readings.

The experimental design of the dataset enables us to further examine the role of question relevant versus question irrelevant information in repeated reading during information seeking. Figure 3 presents TF and SR in first and repeated reading in Hunting, split by words that are inside the critical span and words that are outside of it. We further distinguish between repeated reading trials in which the critical span is the same across the two readings and those in which they differ. First, we observe reading facilitation both inside and outside the critical span ($p < 10^{-73}$ inside and outside the critical span for both TF and SR), with a larger effect outside the critical span ($p < 10^{-5}$ for both TF and SR)[6]. This result is partially at odds with Kaakinen & Hyönä (2007), who found repeated reading facilitation for both relevant and irrelevant text, but found no interaction between repeated reading and the relevance of the textual span to the task. For control, Figure 6 in the SM presents this analysis in Gathering, with no differences between within and outside the critical span.

Figure 3, and SM Figure 4 and Figure 5, show that facilitation effects in repeated reading are larger when the first and the repeated readings share the same critical span for all measures ($p < 0.05$ for all) except for first pass Gaze Duration ($p = 0.32$)[7]. This overall difference is driven by lower repeated reading times within the critical span when the first and the repeated readings share the same critical span ($p < 10^{-3}$ for all measures except for first pass Gaze Duration where $p = 0.12$), while differences outside the critical span are not significant for all measures. The interaction of same critical span – inside the critical span is significant in TF, SR, Regression Rate and fixation count ($p < 0.05$ for all four)[8]. Overall, these results indicate that reading facilitation during repeated reading in information seeking occurs both inside and outside task critical spans. It is greater outside such segments, and further depends on whether the distribution of task critical and non-critical information is identical across readings. More broadly, both the relevance of specific text segments to the task and the similarity between the tasks with respect to these segments across the two readings play substantial roles in reading facilitation during information seeking.

Response to Linguistic Word Properties

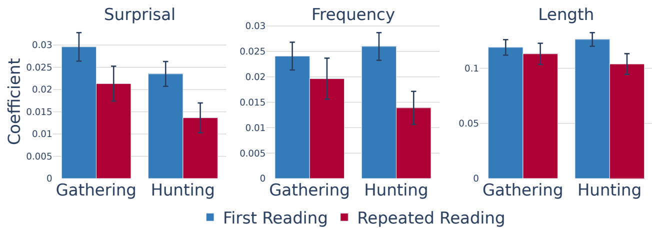
Here, we analyze the strength of the behavioral response to three linguistic word properties which have been shown to be robust predictors of reading times: predictability, frequency and length. We quantify predictability using surprisal (Hale, 2001; Levy, 2008), defined as $-\log_2(p(word|context))$, where *context* is the textual content preceding the *word* in the given paragraph. We extract surprisal values from GPT-2 (Radford et al., 2019), a language model with good correlations with reading times (Wilcox et al., 2020; Shain et al., 2022). For frequency estimates, we compute unigram surprisal, $-\log_2(p(word))$ using frequency counts from Wordfreq (Speer, 2022), which is based on multiple corpora. Following common practice, in this analysis we exclude words that start or finish a line, numbers, and words with punctuation.

Figure 4 presents the current word coefficients from a linear mixed effects model that predicts TF from these properties

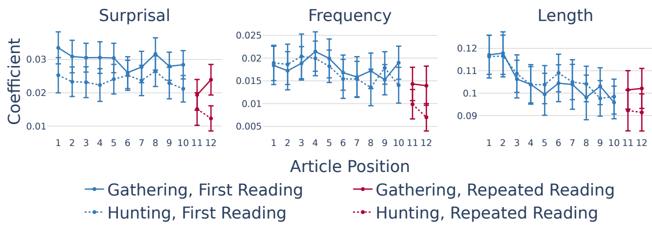
[6] Interaction term $repeated : in\_cs$ in $RT \sim repeated * in\_cs + (repeated * in\_cs|subj) + (repeated * in\_cs|parag)$.

[7] $same\_cs$ term in $RT \sim same\_cs + (same\_cs|subj) + (same\_cs|parag)$ in repeated reading only.

[8] $same\_cs : in\_cs$ term in $RT \sim same\_cs * in\_cs + (same\_cs * in\_cs|subj) + (same\_cs * in\_cs|parag)$ in repeated reading only.



(a)



(b)

Figure 4: The effects of surprisal, frequency, and word length on Total Fixation Duration (TF). TF times are z-normalized separately for each combination of reading regime (Hunting, Gathering) and text presentation (first, repeated). Depicted are current word coefficients from linear mixed-effects models that predict TF times from these properties of the current and previous words[9]. Error bars represent 95% confidence intervals. (a) Overall mean coefficients (b) Mean coefficients as a function of article position in the experiment. Positions 1–10: first reading. Position 11: a consecutive repeated reading of the article in position 10. Position 12: a non-consecutive repeated reading of one of the articles 1–9.

of the current and previous words[9]. In this analysis we z-normalize TF times in each combination of reading regime (Hunting, Gathering) and text presentation (first, repeated) to control for the differences in reading speed across these conditions observed in Figure 1. Across the two reading regimes, repeated reading is marked by diminished responsiveness to linguistic properties of words[10] ($p < 10^{-5}$ for all properties in both regimes) with the exception of the length effect difference in Gathering which is not significant ($p = 0.66$). In line with our analysis in Figure 1a, we find larger reductions of frequency and length effects in repeated reading in Hunting compared to Gathering ($p < 0.03$ for both properties, see formula in Footnote 1 of the SM). Figure 8a and Figure 9a in the SM depict this analysis for First Fixation and Gaze duration, with similar results. Figure 7 in the SM presents this analysis with raw TF, where differences in word property effects across first and repeated reading persist, and further hold for the length effect difference in Gathering.

In Figure 4b we break down these results by the position of the article in the experiment. Similarly to the reading times analysis in Figure 1b, we observe that all three effects diminish with article position during the first reading in both reading regimes ($p < 10^{-4}$ for all three effects), except for the surprisal effect in Hunting and frequency effect in Gathering. Further, in line with our global reading time measures results, we find increasing surprisal effect across positions 11 and 12 in Gathering ($p < 10^{-3}$). For the rest of the effect differences across positions 11 and 12 the trends are similar, but not significant. Figure 8b and Figure 9b in the SM present this analysis for First Fixation and Gaze Duration, with similar results regarding the differences across positions 11 and 12.

In Figure 10 in the SM we present the effects of word properties on z-normalized TF, FF and GD in article 12, as a function of its position during the first reading. Similarly to TF in Figure 2 we observe no difference across positions in both Hunting and Gathering. Overall, we find that trends observed for global fixation measures largely extend to the response to linguistic word properties.

Individual Differences

Our final analysis focuses on individual differences in repeated reading effects. We examine two key axes of variation across readers which could a-priori modulate such effects, reading speed and reading proficiency. We quantify reading proficiency using the reader's reading comprehension performance during the first reading. We hypothesise that slower readers will have larger repeated reading facilitation effects due to increased *quantity* of interaction with the text during the first reading. We similarly predict larger facilitation effects for more proficient readers due to higher *quality* of the interaction with the text during the first reading. Reading proficiency was previously shown to interact with repeated reading effects in children (Zawoyski et al., 2015).

Figure 5 depicts the difference in TF and SR across the two readings as a function of the participant's reading speed and comprehension accuracy during the first reading. We find a significant negative correlation with reading speed for TF in both Hunting ($R^2 = 0.52, p < 10^{-20}$) and Gathering ($R^2 = 0.45, p < 10^{-25}$) and SR in Hunting ($R^2 = 0.07, p < 10^{-4}$). We further observe that the reading speed effect is larger in Hunting compared to Gathering in both TF ($p < 0.05$) and SR ($p < 10^{-3}$)[11]. Similar results for all other measures are presented in SM Figure 11a. SM Figure 11b shows that these results also hold when examining the differences in the raw TF response to surprisal ($p < 10^{-13}$), frequency ($p < 0.02$) and length ($p < 10^{-5}$) in both regimes, with no significant interactions between reading regime and effect difference.

Differently from the results obtained for reading speed, and contrary to our hypothesis, we find no correlation between reading comprehension performance and repeated reading

[9] $TF \sim freq * len + surp + freq_{prev} + len_{prev} + surp_{prev} + (freq + len + surp|subj) + (freq + len + surp|parag)$ fitted separately to each combination of reading regime and text presentation. *prev* refers to the previous word, to account for spillover effects (Rayner, 1998).

[10] Interaction terms with *repeated* in $TF \sim freq : len + freq * repeated + len * repeated + surp * repeated + freq_{prev} + surp_{prev} + len_{prev} + (repeated + freq + surp + len|subj) + (repeated + freq + surp + len|parag)$ applied to each reading regime.

[11] Interaction term $regime : speed$ in $\Delta \sim regime * speed$

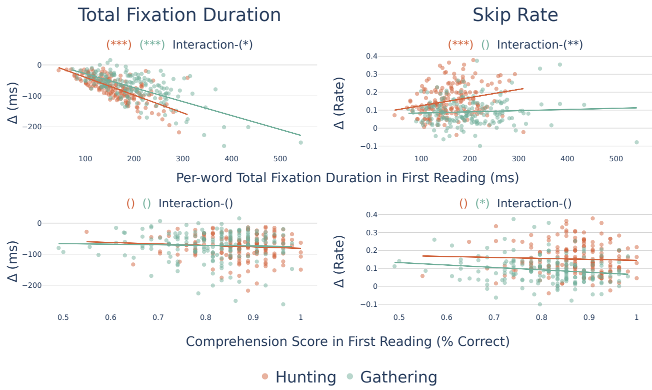


Figure 5: Individual differences in repeated reading effects. The graphs depict the second minus the first reading Total Fixation (TF) and First Pass Skip Rate as a function of participants' reading speed measured with mean per word TF in the first reading (top), and the percentage of reading comprehension questions answered correctly during the first reading (bottom). Each circle represents a participant. We consider only the subset of articles that were read twice by each participant. Above each graph is the significance of the two slopes and the interaction between them. '(***)' p < 0.001, '(**)' p < 0.01, '(*)' p < 0.05, '()' p > 0.05.

facilitation. In the SM, we show that this also holds for all other examined measures (Figure 12a) and the word property effects (Figure 12b). We note that this is the case despite larger improvements in reading comprehension performance during repeated reading for participants with lower comprehension performance during first reading ($p < 10^{-7}$ in Hunting, $p < 10^{-3}$ in Gathering, Figure 12c in the SM).

Discussion

Our work provides a detailed investigation of eye movements in repeated reading. We find that repeated reading is characterized by robust reading facilitation, as exhibited both by global eye movement measures and the response to linguistic text properties. Our analyses further reveal that these effects can be modulated by when and how the first text was read. In ordinary reading, repeated reading benefits are attenuated in the presence of intervening material between readings. In information seeking, the effects are larger for task irrelevant information, and also larger when the division to task relevant and irrelevant information is identical across the two readings. We also find that reading facilitation is modulated by the reading speed of the participant and not by their reading proficiency. Overall, we tend to find larger facilitation effects and individual differences in information seeking compared to ordinary reading.

Our study has a number of limitations. First, the experimental setup is limited to the duration of a single one-hour long eye-tracking session. Further, within this framework, consecutive repeated reading does not constitute immediate repeated reading that involves working memory. Between the two readings of a paragraph in articles 10 and 11 there are 3-6 intervening paragraphs, with a reading comprehension question after each paragraph. We leave the investigation of both shorter and longer repetition intervals to future work. Additionally, although we examine two different reading regimes, their coverage is limited. We use the term ordinary reading to refer to reading without a specific reading goal beyond general text comprehension. Following Huettig & Ferreira (2022), we acknowledge that the term itself is not without faults and that the experimental setup used here is not fully representative of such reading modes. It is also not entirely naturalistic, especially with respect to the presence of a reading comprehension question after each paragraph. In information seeking, while question answering is a general framework for formulating information seeking tasks, the experiment captures only a limited range of tasks. In particular, the search scope is restricted to a single paragraph, the tasks are constrained by the annotation guidelines imposed by OneStopQA, and readers who forgot the task cannot return to it during paragraph reading. Future work with different tasks, amounts of text, and experimental setups is needed.

Despite these limitations, our results provide a number of key insights. In line with prior work, they suggest that repeated reading is much more efficient than first reading. Prior exposure to a text at any stage of an hour-long reading session leads to drastic reductions in reading times and attenuated responses to linguistic word properties. Interestingly, while we find an effect of intervening textual material, it is restricted to ordinary reading, relatively modest, and is not sensitive to the precise amount of intervening material. This suggests robust and durable encoding of the text during the time frame of the experiment, and potential encoding differences during ordinary reading and information seeking. We further find that in information seeking, the effectiveness of this encoding depends on task similarity across readings. Our individual differences analysis suggests that it is the amount of prior processing of the text rather than its effectiveness that impacts repeated reading facilitation. Taken together, these results are in line with a resource-rational approach to language processing, whereby prior processing of the text enables readers to greatly reduce the cognitive effort invested in repeated reading, within constraints that are imposed by memory and by the relevance of the first interaction with the text to the requirements of the second reading.

Conclusion

In this work, we compared eye movements in first and repeated reading, examining repeated reading facilitation, the effect of intervening material, the relevance of specific text segments to the task and variation across readers. Our analyses challenge and extend prior literature and provide a detailed empirical picture that characterizes eye movements in repeated reading in a fine-grained manner. These results can inform future models of eye movements in reading that will account not only for first reading, but also for repeated reading.

Acknowledgments

This work was supported by ISF grant 2070731.

References

Ardoin, S. P., Eckert, T. L., & Cole, C. A. (2008). Promoting generalization of reading: A comparison of two fluency-based interventions for improving general education student's oral reading rate. *Journal of Behavioral Education*, *17*, 237–252.

Berzak, Y., Malmaud, J., & Levy, R. (2020). STARC: Structured Annotations for Reading Comprehension. *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics.*.

Faulkner, H. J., & Levy, B. A. (1999). Fluent and nonfluent forms of transfer in reading: Words and their message. *Psychonomic Bulletin & Review*, *6*, 111–116.

Foster, T. E., Ardoin, S. P., & Binder, K. S. (2013). Underlying changes in repeated reading: An eye movement study. *School Psychology Review*, *42*(2), 140–156.

Hahn, M., & Keller, F. (2023). Modeling task effects in human reading with neural network-based attention. *Cognition*, *230*, 105289.

Hale, J. (2001). A probabilistic earley parser as a psycholinguistic model. In *Second meeting of the north american chapter of the association for computational linguistics.*

Huettig, F., & Ferreira, F. (2022). The myth of normal reading. *Perspectives on Psychological Science*.

Hyönä, J. (1995). An eye movement analysis of topic-shift effect during repeated reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*(5), 1365.

Hyönä, J., & Niemi, P. (1990). Eye movements during repeated reading of a text. *Acta psychologica*, *73*(3), 259–280.

Kaakinen, J. K., & Hyönä, J. (2007). Perspective effects in repeated reading: An eye movement study. *Memory & cognition*, *35*, 1323–1336.

Karpicke, J. D., Butler, A. C., & Roediger III, H. L. (2009). Metacognitive strategies in student learning: Do students practise retrieval when they study on their own? *Memory*, *17*(4), 471–479.

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

Kuhn, M. (2004). Helping students become accurate, expressive readers: Fluency instruction for small groups. *The Reading Teacher*, *58*(4), 338–344.

Levy, R. (2008). Expectation-based syntactic comprehension. *Cognition*, *106*(3), 1126–1177.

Malmaud, J., Levy, R., & Berzak, Y. (2020). Bridging Information-Seeking Human Gaze and Machine Reading Comprehension. In *Proceedings of the 24th Conference on Computational Natural Language Learning* (pp. 142–152). Stroudsburg, PA, USA: Association for Computational Linguistics.

Meyer, M. S., & Felton, R. H. (1999). Repeated reading to enhance fluency: Old approaches and new directions. *Annals of dyslexia*, *49*, 283–306.

Radford, A., Wu, J., Child, R., Luan, D., Amodei, D., & Sutskever, I. (2019). Language models are unsupervised multitask learners. *OpenAI blog*, *1*(8), 9.

Raney, G. E., & Rayner, K. (1995). Word frequency effects and eye movements during two readings of a text. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale*, *49*(2), 151.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological bulletin*, *124*(3), 372.

Rayner, K., Slattery, T. J., Drieghe, D., & Liversedge, S. P. (2011). Eye movements and word skipping during reading: Effects of word length and predictability. *Journal of Experimental Psychology: Human Perception and Performance*, *37*(2), 514–528. (Place: US Publisher: American Psychological Association)

Rayner, K., Warren, T., Juhasz, B., & Liversedge, S. (2004, December). The Effect of Plausibility on Eye Movements in Reading. *Journal of experimental psychology. Learning, memory, and cognition*, *30*, 1290–301.

Shain, C., Meister, C., Pimentel, T., Cotterell, R., & Levy, R. P. (2022, Nov). *Large-scale evidence for logarithmic effects of word predictability on reading time.* PsyArXiv.

Shubi, O., & Berzak, Y. (2023). Eye movements in information-seeking reading. In *Proceedings of the annual meeting of the cognitive science society* (Vol. 45).

Speer, R. (2022, September). *rspeer/wordfreq: v3.0.* Zenodo.

Teigen, T., Malanga, P. R., & Sweeney, W. J. (2001). Combining repeated readings and error correction to improve reading fluency. *Journal of Precision Teaching and Celeration*, *17*(2), 58–67.

Wilcox, E. G., Gauthier, J., Hu, J., Qian, P., & Levy, R. (2020). On the predictive power of neural language models for human real-time comprehension behavior. *arXiv preprint arXiv:2006.01912*.

Zawoyski, A. M., Ardoin, S. P., & Binder, K. S. (2015). Using eye tracking to observe differential effects of repeated readings for second-grade students as a function of achievement level. *Reading Research Quarterly*, *50*(2), 171–184.