Eye Movements in Repeated Reading

Yoav Meiri, Yevgeni Berzak ; Technion - Israel Institute of Technology

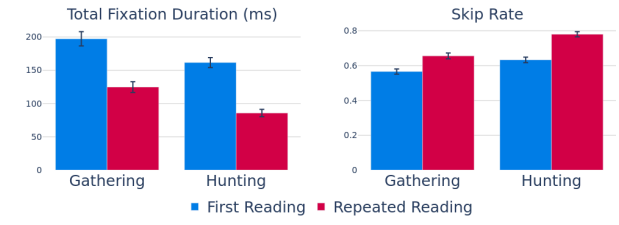
Introduction. In this work, we study how repeated reading affects reading times and their sensitivity to linguistic word properties. We examine these questions across two different reading regimes, ordinary reading and information-seeking. Consistent with prior findings [1-2, among others], we observe that global eye movement measures reflect reading facilitation in repeated reading. Differently from [2], we find that repeated reading leads to a smaller frequency effect, and further find smaller effects for surprisal and word length. Finally, we find that repeated reading facilitation can be modulated by *when* the text was first presented, as well as by *how* it was processed across the two readings.

Design. We use OneStop [3], an English eyetracking dataset with 360 participants and 30 articles with 162 paragraphs from the Guardian. Each paragraph has 3 multiple choice reading comprehension questions. The textual segment essential for answering a question correctly, called the critical span (CS), is manually annotated in the paragraph. An experimental trial consists of reading a paragraph, followed by answering a question on a new screen. Each participant reads 10 articles, followed by two articles that are presented for a second time with identical text but a different question for each paragraph. The 11th article is a second presentation of the article in position 10, and the 12th article is a second presentation of one of the articles in positions 1-9. 180 participants are in an information-seeking regime (Hunting) in which the question (but not the answers) is presented prior to reading the paragraph, and 180 are in an ordinary reading regime (Gathering) where they see the question only after the paragraph. There are 1,723,983 data points during first reading, and 344,774 during repeated reading. Mean experiment duration is 57 minutes in Gathering and 54 minutes in Hunting.

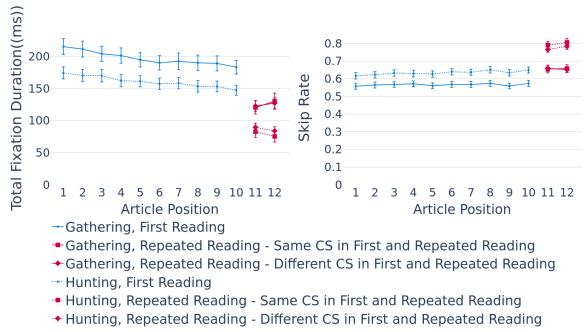
Results. In Fig. 1a, we examine Total Fixation duration (TF) and Skip Rate (SR), and find that in repeated reading TF sharply decreases and SR sharply increases in both reading regimes ($p < 10^{-59}$). Fig. 1b breaks down these results as a function of article position in the experiment. During the first reading, TF decreases ($p < 10^{-25}$) and SR increases ($p < 10^{-4}$) with article position. Contrary to this trend, in Gathering, in position 12 TF increases ($p < 0.04$). In Hunting, we find greater rereading facilitation when the first and second readings have the same critical span (for TF, SR $p < 0.04$). In Fig. 2a we examine the reading time response to word properties using the coefficients of mixed-effects models that predict TF from frequency, surprisal (GPT-2 [4]) and word length. In all cases the effects are smaller in repeated reading ($p < 10^{-26}$). Fig. 2b is analogous to Fig. 1b, but traces word property effects. The trends observed for global fixation measures are reflected in part in these effects. All three effects diminish with position in first reading ($p < 10^{-3}$). In Gathering, the surprisal effect increases between articles 11 and 12 ($p < 10^{-3}$). In Hunting, the frequency effect in repeated reading is smaller when the first and second readings share the critical span ($p < 10^{-3}$).

Discussion. We show large reading facilitation effects across ordinary reading and information-seeking in repeated reading. These effects are exhibited both in global fixation measures, and in the response to linguistic word properties. We further find that in ordinary reading, facilitation is reduced in the presence of intervening textual material between the two readings, an effect that can be potentially attributed to memory decay over time. The absence of this effect in information-seeking may be related to shorter time spans between readings, different processing strategies, and potentially more efficient information encoding in this regime. We leave a comprehensive analysis and interpretation of this difference to future work. In information-seeking, we find larger facilitation effects when task relevant and task irrelevant information is identical across readings. Our work provides a detailed empirical account of repeated reading, which can inform psycholinguistic theory and future models of eye movements in reading that will take into account repeated reading in different reading regimes.

References. **[1]** Hyönä, J., & Niemi, P. (1990). Eye movements during repeated reading of a text. *Acta Psychologica*. **[2]** Raney, G. E., & Rayner, K. (1995). Word frequency effects and eye movements during two readings of a text. Canadian Journal of Experimental Psychology. **[3]** Malmaud, J., Levy, R. P., and Berzak, Y., (2020) Bridging Information-Seeking Human Gaze and Machine Reading Comprehension. *Proc. CoNLL*. **[4]** Radford, A, et al. (2019) Language models are unsupervised multitask learners.

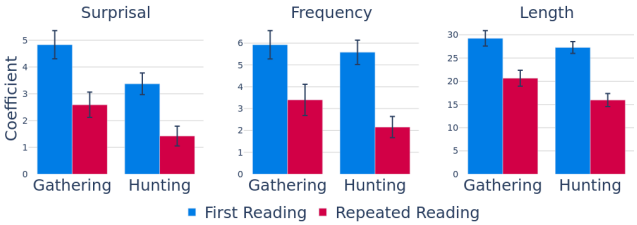


(a)

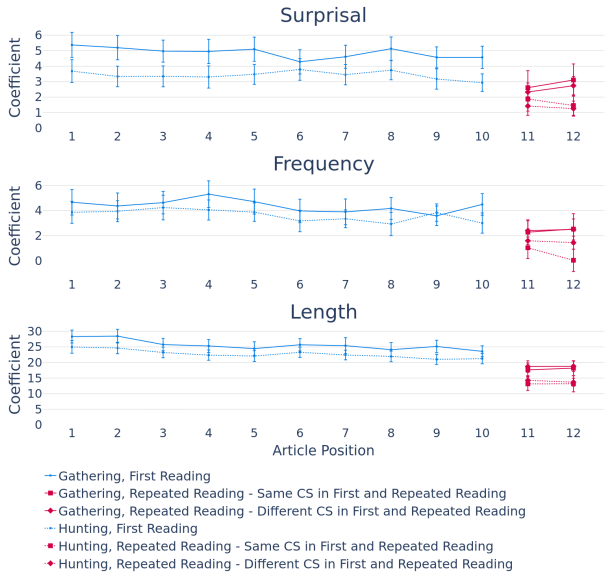


(b)

Figure 1: Mean per word Total Fixation Duration and Skip Rate with 95% confidence intervals across ordinary reading (Gathering) and information seeking (Hunting) in first and repeated reading. (a) Overall means (b) Means as a function of the article presentation position in the experiment. Positions 1–10: first reading, position 11: a consecutive rereading of the article in position 10, position 12: a non-consecutive rereading of one of the articles 1–9. In positions 11 and 12 we further distinguish between trials in which the questions in the first and repeated readings have the same critical span, and trials in which they have distinct critical spans.



(a)



(b)

Figure 2: The effects of surprisal, frequency and length on Total Fixation Duration. Depicted are current word coefficients from linear mixed-effects models predicting TF times from these properties of the current and previous words fitted separately for each reading regime (Hunting, Gathering) and text presentation mode (First, Repeated). Error bars represent 95% confidence intervals. (a) Overall means (b) Means as a function of article position in the experiment. Positions 1–10: first reading, position 11: a consecutive rereading of the article in position 10, position 12: a non-consecutive rereading of one of the articles 1–9. In positions 11 and 12 we further distinguish between trials in which the questions in the first and repeated readings have the same critical span, and trials in which they have distinct critical spans.