Effects of linguistic context and reading abilities on processing of unknown words

Margarita Ryzhova (mryzhova@lst.uni-saarland.de)[1], Emilia Ellsiepen[2] & Vera Demberg[3]
[1,2,3]Department of Language Science and Technology, Saarland University, Germany

Abstract

Encountering unknown words disrupts reading and requires readers to rely on context to infer meaning. We investigate how contextual expectations (science-themed vs. daily-life narratives) and reader literacy modulate processing of unknown words. Using pseudowords as controlled proxies for unfamiliar lexical items, we conducted an eye-tracking experiment (Experiment 1) and a self-paced reading study (Experiment 2). To investigate how participants infer the meaning of pseudowords, each pseudoword was followed by a semantically associated word in the subsequent sentence. Across both methods, pseudowords elicited longer reading times. Crucially, eye-tracking results revealed that higher-literacy readers showed greater sensitivity to the semantically associated word, particularly in science-themed contexts. While both groups of readers showed attempts at reconstruct meaning in an offline task, higher-literacy readers mentioned unknown words themselves more frequently, which can be taken as evidence for incidental vocabulary learning. Together, these findings suggest that literacy supports strategic, context-sensitive processing of unknown words.

Keywords: reading; eye-tracking; literacy; text genres; individual differences; novel words; incidental word learning

Introduction

Encountering unknown words can be a source of processing difficulty during reading. Although such words are particularly frequent in domain-specific and technical texts, they are not limited to specialized literature and may also appear unexpectedly in otherwise familiar contexts. From a broader perspective, reading itself is a primary mechanism through which readers encounter new and unfamiliar words across a wide range of contexts. Consistent with this, reading has been shown to increase vocabulary size, suggesting that incidental word learning through repeated exposure to unknown words in meaningful contexts is a common process during reading (Stanovich et al., 1998). Experimental work on online processing has shown that readers slow down on unfamiliar words, but are able to infer some meaning from the sentence context even after single exposure (Williams & Morris, 2004). The role of the broader, text-level context in shaping these processes, however, is underexplored (see e.g., Wochna & Juhasz, 2013). Our study investigates how the type of context in which unknown words appear influences the reading patterns of unknown words by shaping different expectations, how semantically associated words are used for inferring the meaning of unknown words, and how readers of different literacy levels differ with regard to reading strategies of texts containing unknown words.

We hypothesize that different types of discourse context may elicit different strategies by shaping readers' expectations about the occurrence of unfamiliar words. In scientific texts, e.g., the reader may expect to encounter specialized terminology that they do not know and thus be less surprised by an unknown word compared to other contexts. Similarly, the topic of a narrative might drive different expectations: In a story set in a hospital or a space station for example, the reader can expect to encounter concepts and words unknown to them more than in a story about a family outing. Following this consideration and to avoid genre effects that might mask our effect of interest, we use short texts that differ in the setting (science-themed vs. daily-life) but share the same narrative style. Unknown words are embedded in these contexts, together with a word that is semantically associated with the intended meaning appearing in the following sentence, allowing us to investigate how semantic cues support meaning reconstruction as reflected in eye-tracking measures and self-paced reading times. Taken together, this design allows us to test not only immediate processing of unknown words but also to detect strategic use of context that may differ between narrative topics.

In addition to population-level effects, substantial evidence indicates that the reading process is influenced by the reader's literacy. More skilled readers are better at monitoring their comprehension, recognizing when additional processing is necessary, and are more motivated to fully understand the texts they read, thereby investing greater cognitive effort to resolve comprehension problems (van den Broek et al., 2009). We hypothesize that strategic use of context to recover the meaning of novel words could also be modulated by literacy.

Finally, we address these questions using two methodologies. Experiment 1 employs eye-tracking to investigate different reading patterns in a more naturalistic setup, while Experiment 2 employs self-paced reading (SPR) to test whether similar effects can be found in a more constrained, but also less expensive reading setup. Although SPR is widely used, it has been criticized for diverging from natural reading. For example, Rayner and Clifton Jr (2002) argue that this divergence may lead readers to adopt different comprehension strategies. This comparison thus allows us to test the extent to which contextual and literacy-related effects generalize across methods.

In G. Boleda, I. Dautriche, M. Maldonado, C. Saldana and T. Torrent (Eds.), *Proceedings of the 48th Annual Conference of the Cognitive Science Society*

Processing Unknown Words

Unknown words have long been known to cause processing difficulty for readers. For instance, Just and Carpenter (1980) studied the reading of scientific texts and found that seven words, that were completely unfamiliar to participants, were processed more slowly than words that participants knew. More recently, these results have been confirmed in analyses of the PoTeC corpus, where unfamiliar technical terms lead to increased processing difficulty (Jakobi et al., 2025; Škrjanec et al., 2023), as reflected in longer fixation durations and more frequent regressions among non-expert readers. The processing of unfamiliar words however also depends on context and expectation: Vigly (2024) showed that the reading times of words which contain clear typos (e.g., *construtcion* instead of *construction*), but are highly expected, hardly lead to a slow-down; instead, reading times increase strongly when encountering unexpected or unknown meaning.

In the experimental setup, it is, however, difficult to control whether participants will know the specific words used in the study. Therefore, a common methodological alternative is to work with pseudowords. A large body of work has validated this method and employed pseudowords to investigate the processing of unknown words in a more controlled manner. Pseudowords preserve orthographic plausibility while eliminating any prior familiarity, ensuring that the target word is unfamiliar to all readers. This makes pseudowords well-suited for examining how readers rely on contextual information and general reading skills when lexical access fails.

Research using pseudowords shows increased reading times not only at the unknown word itself but also at informative constraining regions that are directly related to the intended meaning (Lowell & Morris, 2014, 2017; Wochna & Juhasz, 2013). Lowell and Morris (2017) further suggest that the observed reading time effects reflect more than mere processing difficulty. Building on earlier findings of Lowell and Morris (2014), they showed that reading times on pseudowords and constraining informative regions correlated with performance on a surprise vocabulary test, suggesting that observed reading patterns index not only failed lexical access, but also potentially word form encoding and meaning assignment. However, since Lowell and Morris (2017) relied on isolated sentences and included a relatively high proportion of unfamiliar items (60 sentences, with half of them containing pseudowords), it remains unclear to what extent the observed effects reflect strategic use of context, learning, or memory-related effects.

Our Study

Our study extends the work of Lowell and Morris (2017) in several regards. Firstly, similar to Wochna and Juhasz (2013), we embed pseudowords (and their real word controls) within short narrative texts, allowing readers to form expectations and integrate information across broader discourse. Secondly, to test meaning reconstruction more directly, we include a semantic associate of the real control word in the sentence following the target word. Finally, we did not administer a vocabulary test, but instead tested for comprehension of the text and meaning assignment for pseudo words in an indirect manner: We asked participants after reading the last text to produce a summary of the same. Our expectation was that participants would only include the pseudo word - either by referring to its meaning, or to the word form - if they already were able to assign meaning to it during reading. This method avoids memory-related effects by not asking about previous items, and offline meaning assignment which might happen if participants are forced to choose a meaning for a word from a given list.

In addition to contextual factors, we focus on individual differences in reading skills. In the present study, we assess readers' literacy using the Author Recognition Test (ART; Stanovich & West, 1989) and a Vocabulary Size Test (Engel & Ettinger, 1997). The ART is widely used as a proxy for print exposure and is based on the assumption that individuals who read more frequently are more likely to recognize the names of authors. Performance on the ART has been shown to correlate moderately to strongly with vocabulary knowledge (Lee et al., 1997; Stanovich et al., 1995).

A substantial body of work demonstrated that reading experience is associated with systematic differences in eye-movement behavior during reading. More experienced readers tend to show more efficient word recognition (Chateau & Jared, 2000; Lowder & Gordon, 2017; Moore & Gordon, 2015), greater use of parafoveal information (Veldre & Andrews, 2014), and more precise saccadic targeting (Kuperman & Van Dyke, 2011). They also typically exhibit reduced effects of word frequency and word length (Kuperman & Van Dyke, 2011).

Crucially, literacy-related differences are not limited to general reading efficiency, but extend to strategic adaptation to text difficulty. For example, Kuperman et al. (2018) show that readers with larger vocabularies adjust their eye-movement behavior as a function of text complexity, exhibiting fewer fixations and higher skipping rates in less complex texts, whereas readers with smaller vocabularies show little modulation across complexity levels. These findings suggest that high-literacy readers are more sensitive to context and more flexible in their reading strategies. We therefore hypothesize that literacy will modulate the processing of unknown words, with higher-literacy readers showing greater sensitivity to contextual information and greater engagement in inferring the meaning of unknown words.

Experiment 1

In Experiment 1, we test whether the context (science-themed vs. daily-life narratives) influences reading behavior of unknown words and associated words, appearing in the next sentence after the novel word was encountered during natural reading. We conducted individual differences tests (ART and vocabulary size tests) to investigate the effect of literacy on this process.

Method

Participants We recruited 96 native German speakers (61.5% female, 1% non-binary, 37.5% male, age range [18, 65], mean = 25.1, sd = 7.2). The experiment lasted around 60 minutes and participants received a compensation of 13 Euros.

Materials 24 German experimental items were constructed. In a critical sentence, a real word noun (target word) that appears approximately in the middle of the sentence was matched with a pseudoword of equal length, generated using the Wuggy tool (Keuleers & Brysbaert, 2010). The sentence following the critical sentence contained a noun that was semantically associated with the target word (associated word). Association strength was not normed, instead the words were chosen to be informative of which type of object the unknown word was, e.g. *precision* in the example indicates that *Finnatte* is most likely a tool. Then these two sentences were placed into two types of short narrative contexts: a daily-life context and a science-themed context. Daily-life stories described familiar, non-specialized situations, whereas science-themed stories described less familiar scenarios involving specialized settings or professions, while retaining a narrative format. Crucially, the critical sentence and the sentence with associate were identical between contexts. An English translation of the example item can be found in Table 1.

This resulted in four versions of each item: daily-life real, daily-life pseudoword, science-themed real, and science-themed pseudoword, yielding a total of 96 experimental stories (24 items $\times$ 2 word conditions $\times$ 2 story context conditions). In addition, 8 filler stories without any word manipulation were constructed for both daily-life and science-themed contexts.

The 96 experimental stories were distributed across 12 experimental lists. Each list contained 16 stories: two stories in each of the four experimental conditions, as well as eight filler stories (four daily-life, four science-themed). With-in lists, each item appeared in only one condition for a given participant.

Each experimental list was further divided into four sublists. The sublists were identical in content, where the order of items was randomized except for the final story. In each sublist, a different pseudoword story from the list appeared as the final item. Thus, across sublists, each pseudoword item served equally often as the final story.

Individual difference measures Literacy levels were assessed with two different tests: a German version of the author recognition test (Grolig et al., 2020) to measure print exposure and a multiple-choice vocabulary test. The vocabulary size test included 55 items of varying difficulty taken from Engel and Ettinger (1997) and Heubeck (2001), see also van Os (2023).

Procedure After signing informed consent, participants were seated in front of an Eyelink 1000. Participants read a single sublist. Each item was presented on a single screen. After reading one item, participants answered a simple comprehension question to ensure focus on the task. Immediately after reading the final story, which always contained a pseudoword, participants were asked to produce a spoken summary of that story as a surprise task. Spoken summaries were audio-recorded and later annotated to assess whether and how the pseudoword was mentioned or interpreted. Following the reading task, participants completed the individual difference tests.

Analysis All data analyses were conducted using the lme4 package in R (Bates et al., 2015; R Core Team, 2024). Post-hoc comparisons were carried out using emmeans package (Lenth & Piaskowski, 2025). We analyzed four different measures on the two regions of interest (critical word, associated word): total reading time (TRT), first-pass reading time (FPRT), regressions in (Reg-In) and regressions out (Reg-Out). Within each region, Bonferroni correction was applied to maintain the nominal alpha level. Continuous outcomes were modeled using gamma mixed effect models with the log link; binary outcomes were analyzed using logistic regression. We included the variables of interest (story condition, word type, literacy) and all interactions as fixed effects. Additionally, we included word length, trial order and the interaction of trial order with word condition as control variables. In the associated region, we furthermore included the normalized lexical frequency of the associated word obtained from DLEX database (Heister et al., 2011). We used the maximal random effect structure licensed by the data, excluding random slopes that resulted in convergence failure or singular fits. All variables were centered. Literacy scores were derived by summing the z-scores of the ART and vocabulary size tests. Prior to computing the literacy composite, ART scores for 12 participants who timed out on more than 40 trials were replaced with the sample mean. The distribution of literacy scores is shown in Figure 1, upper row.

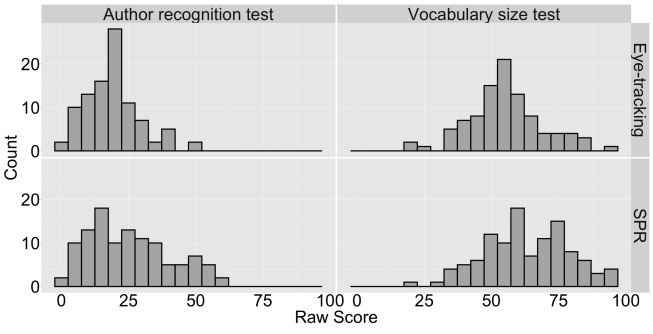

Figure 1: Distribution of raw ART and Vocabulary scores across Experiment 1 (eye-tracking; upper row) and Experiment 2 (SPR; lower row). Higher scores indicate better performance on the respective tests.

Table 1: Example item translated to English with critical and post-critical sentence in German with English glosses. Critical word and associate are in bold. Line breaks within the story text were not shown to participants. Spacing for critical and post critical sentence corresponds to chunks in Experiment 2 (SPR).

	daily-life	science-themed
pre-critical context	Little Adam sat on the floor of his nursery and was eagerly building a house out of Lego bricks. When he wanted to insert a small brick, it slipped from his fingers and disappeared into the Lego house. Adam began to cry, whereupon his older brother Max, who was playing in the next room rushed over.	In the emergency room of a hospital there was bustling activity. Doctors and nurses hurried from one patient room to another. Around midnight a patient was brought in who seemed to be suffering from severe pain. Dr. Belau took over the operation as chief surgeon.
critical sentence	Er liess sich sofort **eine Finnatte / eine Pinzette** reichen, nachdem er das Problem erkannt hatte. He let himself immediately **prodents / tweezers** hand, after he the problem recognized had. *He immediately had a pair of prodents / tweezers handed to him after recognizing the problem.*	
associate	Die Arbeit erforderte **die Präzision** eines Experten. The work required **the precision** of an expert.	
post-critical context	Adam looked over his brother's shoulder while he searched for the brick. Finally, he held the tiny building block in his hand and gave it to Adam, who beamed at his older brother. Shortly after, the house was finally finished. Now Adam had to build a car for the Lego figure.	The operation was successful, and the patient could be discharged again after a short time. Dr. Belau worked until five o'clock in the morning before he collapsed into bed, exhausted. He knew that he had to prepare for the next day of work, because the well-being of the patients was his top priority.

Results

Target word region In the target region, we consistently observed effects of word condition across all eye-tracking measures, with pseudowords eliciting longer reading times and more regressions than real words.

In the model of FPRT, we additionally observed a significant interaction between literacy and story context ($est = 0.06$, $SE = 0.02$, $t = 2.93$, $p_{adjusted} = .01$), see Figure 2. Pairwise comparisons of estimated marginal means showed that for high-literacy readers ($+1SD$), reading times differed significantly between daily-life and science-themed contexts, with longer reading times in the science-themed condition ($est = -0.14$, $SE = 0.06$, $z = -2.41$, $p < .05$). In contrast, no reliable difference between story contexts was observed for low-literacy readers ($-1SD$; $est = 0.10$, $SE = 0.06$, $z = 1.74$, $p = .083$). This pattern indicates that early sensitivity to story context during lexical processing is present for high-literacy readers, whereas low-literacy readers show little differentiation between contexts at the target word.

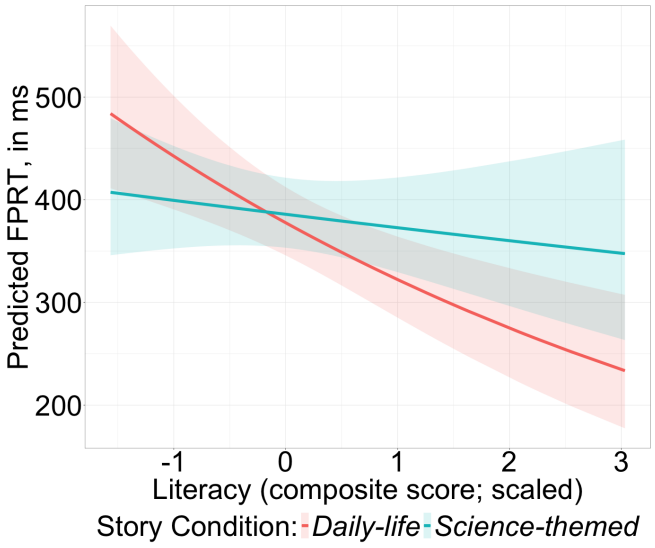

Figure 2: Target word region. Predicted first-pass reading time (FPRT) as a function of literacy (standardized composite score) and story context. Shaded areas indicate 95% confidence intervals.

Associate word region In the model of TRT, there was a significant 3-way interaction between literacy, word, and context ($est = 0.07$, $SE = 0.02$, $t = 3.63$, $p_{adjusted} = .001$). A post-hoc analysis revealed that high-literacy readers exhibited significantly longer reading times following pseudowords than real words in science-themed contexts ($est = -0.25$, $SE = 0.10$, $p < .01$), whereas no such differences were observed in daily-life contexts. Low-literacy readers showed no significant pseudo/real word differences in either context. This is also clearly visible in Figure 3, where readers with higher literacy are predicted to show longer reading times than low literacy readers only for the associate following a pseudoword in science-themed context, whereas for all other combinations, there is a speed-up.

Summary production Due to technical issues, spoken summary data were missing for 12 participants. The remaining 84 answers were analyzed and categorized with regard to whether the pseudoword was mentioned at all (39.3% of all trials), and in what form: some participants repeated the pseudoword themselves, the majority picked up on the meaning, and in 3 instances the control word was mentioned. Table 2 shows

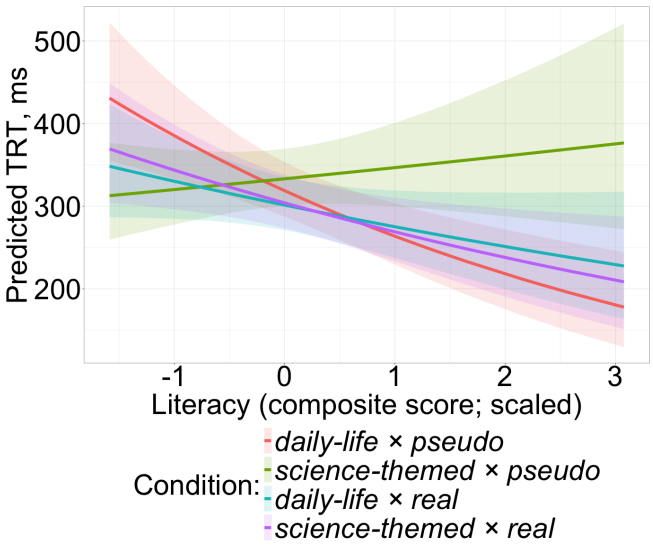

Figure 3: Semantic associate region. Predicted total reading time (TRT) as a function of literacy (standardized composite score), word, and story context. Shaded areas indicate 95% confidence intervals.

the distribution by literacy level (median split): While there is almost no difference between the likelihood to mention it or not, high literacy readers show an increased tendency to produce the actual pseudoword. This can be characterized as incidental vocabulary learning, while low literacy readers focused more on understanding the story. Fisher's exact test confirmed a significant association between literacy level and type of mentioning (p = .025). Due to sparseness of data, we were not able to perform more advanced modeling including random effects. There were no notable differences in distribution between different contexts.

Table 2: Frequencies of mentioning the pseudoword in different forms within the summary production task by literacy level (median split).

	High literacy	Low literacy
pseudoword	7	2
meaning	6	15
control word	2	1

Experiment 2

In Experiment 2, we tested the same hypotheses as above during self-paced reading (SPR). While SPR is an efficient way to conduct reading research, it is unclear whether different reading patterns can be captured to a similar degree as with eye-tracking.

Method

Participants We recruited 111 participants (39.6% female, 60.4% male, age range [19, 71], mean = 33.5, sd = 10.8) on Prolific. The experiment lasted around 60 minutes and was fully online.

Materials Materials were identical to Experiment 1. The stories were chunked: while most chunks contained 2-5 words, the target region always included only the critical word with its determiner. The spill-over region (the chunk immediately following the target region) was controlled across items to contain either one longer content word (e.g., *reichen* in Table 1) or two words – a short function word followed by a content word. The associated word region also included the semantic associate word with its determiner.

Procedure Similar to Experiment 1, participants first performed the reading task and subsequently the individual difference tasks. For the reading task, texts were presented chunk by chunk in the moving window self-paced reading paradigm. There was no summary task in Experiment 2.

Analysis Reading times were analyzed in the target word region (real vs. pseudoword) and the subsequent spill-over region, as well as at the associated word. The data were analyzed following the same procedure as described in Experiment 1. Each model included the following control variables: number of characters, chunk order, trial order, and interaction of trial order with word condition as controls. Additionally the models for spill-over and associated word regions included normalized lexical frequency of the content word.

Two trials were excluded due to extremely short (12 ms) and extremely long (12000 ms) reading times. Prior to computing the literacy composite, ART scores for 3 participants were replaced with the sample mean due to timeouts, as described in Experiment 1.

Results

We observed significant effects of word condition in both the target and spillover regions (target: est = 0.16, t = 4.29, $p_{adjusted}$ < .001; spill-over: est = 0.17, t = 5.60, $p_{adjusted}$ < .001), with pseudowords eliciting longer reading times than real words. This pattern is consistent with the results of Experiment 1. In addition, the model for the target region included a significant interaction between trial order and word condition (est = −0.03, t = −2.47, $p_{adjusted}$ = .02), indicating that pseudowords were read faster over the course of the experiment. In the spill-over region, we additionally observed a main effect of story context, with science-themed stories read faster than daily-life stories (est = −0.02, t = −2.25, $p_{adjusted}$ = .04).

In contrast to the eye-tracking results, the self-paced reading experiment did not show evidence for the literacy × story interaction (target: est = −0.02, t = −1.22, $p_{adjusted}$ = .44; spill-over: est = −0.01, t = −0.91, $p_{adjusted}$ = .72),

nor a three-way interaction of literacy × word × story in the associated region ($est = -0.01$, $t = -0.96$, $p_{adjusted} = .68$).

General Discussion

Our experiments show that pseudowords (as a proxy for novel words) result in increased processing time throughout: we see an increase in early (first-pass reading time) and late (total reading time) eye tracking measures, as well as more regressions in and out in Experiment 1, and increased reading times on the target and the spill-over region in self-paced reading. We did not find interactions between context type and word type, so on average the hypothesized context-driven expectations did not influence reading behavior of the critical word or the associate. Interestingly, we found effects of literacy in some eye-tracking measures: for high literacy readers, we found longer reading times in the science-themed context on the target region irrespective of word type. On the associate word, on the other hand, there was a 3-way interaction driven by longer reading times in the pseudoword condition in science-themed contexts for high literacy readers only. We did not see an effect of literacy in Experiment 2 using SPR. In our offline measure based on a surprise summary production after Experiment 1, we saw a substantial number of summaries that referred to the pseudoword (either by mentioning it explicitly or by referring to its meaning) for readers of all literacy levels and both contexts, but a difference between the literacy groups in whether they showed a greater tendency to produce the word itself, or the meaning.

For high literacy readers, we find three interesting effects: the general slow down in the target region in science-themed context as opposed to daily-life context suggests strategic adaptation. Importantly, contexts were not constructed to be of different complexity, avoiding technical terms and keeping mean overall surprisal at a comparable level across story contexts. Possibly, high-literacy readers were thus expecting more difficult words and adjusted their general reading behavior accordingly. Alternatively, this effect can be characterized as a missing speed-up that we see in the daily-life context, in line with findings by Kuperman et al. (2018). The more careful reading of science-themed narratives, in turn, might have driven the second effect: Increased attention on the associated word for high literacy readers. We propose that the general tendency to approach science-themed narratives in a different manner enabled them to identify the associate as relevant for the meaning reconstruction of the novel word. This specific behavior, however, was apparently not a prerequisite of meaning reconstruction, as mentions of the pseudoword's underlying concept did not differ between literacy groups in post-experiment summaries. The third effect, the higher tendency to reproduce the pseudoword itself offline suggests increased ability of incidental word learning, although this insight is based on a comparatively small number of instances and should be interpreted cautiously. Our findings extend Joseph and Nation (2018), who found that children's reading times of novel words were predicted by their comprehension scores, but only after encountering them multiple times.

Although general processing difficulty was observed in both, eye-tracking and SPR, reader-specific strategies were only observed in eye-tracking and not in SPR. There are two possible explanations, one is the difference between the two *methods*, while the other is concerned with differences between the two *samples*. A difference between the methods might be observed, because different reading strategies involve differences in skipping rates or strategic regressions (Hyönä et al., 2002; Kuperman et al., 2018), both of which are not possible in SPR. It seems thus that forcing participants into linear reading interferes with their natural reading behavior and obscures individual differences, if they are related to different reading strategies. Additionally, we saw an interaction between trial number and word type with a stronger speed-up for pseudowords in SPR. This might point to a reduced level of processing depth as participants could not revisit important parts of the context to be able to reconstruct the meaning. The alternative explanation focuses on the differences between our two samples: as can be seen in Figure 1, we observe more symmetrical distributions of ART and vocabulary size for the SPR sample, with more high values. The reason for this difference is likely connected to the different age distribution, as the in-lab eye-tracking study mainly recruited students, while the available prolific population is more diverse and ART as well as vocabulary size in general increases with age (Stanovich et al., 1995). While age might have influenced general reading speed, the wider distribution in ART and vocabulary size, should have given us **more** power to detect an effect of literacy in SPR, which we did not observe. For this reason, the task difference is a more likely cause for the discrepancy between eye-tracking and SPR results.

Limitations and future work

Differences between contexts were only seen for the group of high literacy readers in our experiments and did not show in all relevant eye tracking measures. We decided on purpose to use a relatively small number of items per participant, because we wanted to record normal reading behavior and prevent participants from focusing on the pseudowords. Our own pilot work (Ryzhova et al., 2023) has shown that reading patterns change substantially after having read multiple pseudowords in the same study. This of course comes at the cost of lower power to detect more subtle effects. It is also possible that our manipulation, science-themed vs daily-life topics in narratives, was not strong enough to elicit more reliable effects and to change the genre, i.e., contrast scientific expository texts and narratives, might make the effects more stable.

Keeping the critical and post-critical sentence constant between contexts allowed a detailed analysis of reading time differences on the same words. In future work, we plan to investigate scan paths of the whole paragraph and extend our analyses to other regions of the text, such as the final word of the critical sentence and the final word of the whole text.

Acknowledgments

This work was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project-ID 232722074 – SFB 1102.

We thank the anonymous reviewers for their helpful comments. We also thank Domenik Mayer and Katharina Trinley for their assistance with materials construction and data collection. We thank Iza Škrjanec for calculating surprisal values and for helpful discussions throughout the development of this work.

References

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Chateau, D., & Jared, D. (2000). Exposure to print and word recognition processes. *Memory & Cognition*, *28*(1), 143–153. https://doi.org/10.3758/BF03211582

Engel, R., & Ettinger, U. (1997). Satzergänzungstest 1. Psychiatrische Klinik der LMU München, Abteilung für experimentelle Psychologie und Psychophysiologie.

Grolig, L., Tiffin-Richards, S. P., & Schroeder, S. (2020). Print exposure across the reading life span. *Reading and Writing*, *33*(6), 1423–1441. https://doi.org/10.1007/s11145-019-10014-3

Heister, J., Würzner, K.-M., Bubenzer, J., Pohl, E., Hanneforth, T., Geyken, A., & Kliegl, R. (2011). Dlexdb–eine lexikalische datenbank für die psychologische und linguistische forschung. *Psychologische Rundschau*.

Heubeck, V. (2001). *Statistische Analyses eines Wortbedeutungstest* [Master's thesis, Universität Konstanz].

Hyönä, J., Lorch Jr, R. F., & Kaakinen, J. K. (2002). Individual differences in reading to summarize expository text: Evidence from eye fixation patterns. *Journal of Educational Psychology*, *94*(1), 44.

Jakobi, D. N., Kern, T., Reich, D. R., Haller, P., & Jäger, L. A. (2025). Potec: A german naturalistic eye-tracking-while-reading corpus. *Behavior Research Methods*, *57*(8), 1–37. https://doi.org/10.3758/s13428-024-02536-8

Joseph, H., & Nation, K. (2018). Examining incidental word learning during reading in children: The role of context. *Journal of Experimental Child Psychology*, *166*, 190–211.

Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological review*, *87*(4), 329.

Keuleers, E., & Brysbaert, M. (2010). Wuggy: A multilingual pseudoword generator. *Behavior research methods*, *42*(3), 627–633.

Kuperman, V., Matsuki, K., & Van Dyke, J. A. (2018). Contributions of reader-and text-level characteristics to eye-movement patterns during passage reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *44*(11), 1687. https://doi.org/10.1037/xlm0000547

Kuperman, V., & Van Dyke, J. A. (2011). Effects of individual differences in verbal skills on eye-movement patterns during sentence reading. *Journal of memory and language*, *65*(1), 42–73. https://doi.org/10.1016/j.jml.2011.03.002

Lee, S.-Y., Krashen, S., & Tse, L. (1997). The author recognition test and vocabulary knowledge: A replication. *Perceptual and Motor Skills*, *85*(3_suppl), 1428–1430.

Lenth, R. V., & Piaskowski, J. (2025). *Emmeans: Estimated marginal means, aka least-squares means* [R package version 2.0.1]. https://rvlenth.github.io/emmeans/

Lowder, M. W., & Gordon, P. C. (2017). Print exposure modulates the effects of repetition priming during sentence reading. *Psychonomic Bulletin & Review*, *24*(6), 1935–1942. https://doi.org/10.3758/s13423-017-1248-1

Lowell, R., & Morris, R. K. (2014). Word length effects on novel words: Evidence from eye movements. *Attention, Perception, & Psychophysics*, *76*(1), 179–189.

Lowell, R., & Morris, R. K. (2017). Impact of contextual constraint on vocabulary acquisition in reading. *Journal of Cognitive Psychology*, *29*(5), 551–569.

Moore, M., & Gordon, P. C. (2015). Reading ability and print exposure: Item response theory analysis of the author recognition test. *Behavior research methods*, *47*(4), 1095–1109. https://doi.org/10.3758/s13428-014-0534-3

R Core Team. (2024). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. Vienna, Austria. https://www.R-project.org/

Rayner, K., & Clifton Jr, C. (2002). Language processing. *Stevens' Handbook of Experimental Psychology*.

Ryzhova, M., Škrjanec, I., Quach, N., Chase, A. V., Ellsiepen, E., & Demberg, V. (2023). Word familiarity classification from a single trial based on eye-movements: A study in German and English. *Proceedings of the 2023 Symposium on Eye Tracking Research and Applications*, Article No. 60. https://doi.org/10.1145/3588015.3590118

Škrjanec, I., Broy, F. Y., & Demberg, V. (2023). Expert-adapted language models improve the fit to reading times. *Procedia Computer Science*, *225*, 3488–3497.

Stanovich, K. E., Cunningham, A. E., & West, R. F. (1998). Literacy experiences and the shaping of cognition. In S. G. Paris & H. M. Wellman (Eds.), *Global prospects for education: Development, culture, and schooling* (pp. 253–288). American Psychological Association. https://doi.org/10.1037/10294-009

Stanovich, K. E., & West, R. F. (1989). Exposure to print and orthographic processing. *Reading research quarterly*, 402–433. https://doi.org/10.2307/747605

Stanovich, K. E., West, R. F., & Harrison, M. R. (1995). Knowledge growth and maintenance across the life span: The role of print exposure. *Developmental Psychology*, *31*(5), 811.

van Os, M. (2023). *Rational speech comprehension: Effects of predictability and background noise* [Doctoral dissertation, Universität des Saarlandes]. https://doi.org/10.22028/D291-40555

van den Broek, P., White, M. J., Kendeou, P., Carlson, S., Wagner, K., Schatschneider, C., & Phythian-Sence, C. (2009). Reading between the lines. *Developmental and individual differences in cognitive processes in reading comprehension. In. K. Wagner, C. Schatschneider, & C. Phythian-Sence (Eds.), Beyond decoding: The behavioral and biological foundations of reading comprehension*, 107–123.

Veldre, A., & Andrews, S. (2014). Lexical quality and eye movements: Individual differences in the perceptual span of skilled adult readers. *Quarterly Journal of Experimental Psychology*, *67*(4), 703–727. https://doi.org/10.1080/17470218.2013.826258

Vigly, J. (2024). *The cost of information: Looking beyond predictability in language processing* [Doctoral dissertation, McGill University].

Williams, R., & Morris, R. (2004). Eye movements, word familiarity, and vocabulary acquisition. *European Journal of Cognitive Psychology*, *16*(1-2), 312–339.

Wochna, K. L., & Juhasz, B. J. (2013). Context length and reading novel words: An eye-movement investigation. *British Journal of Psychology*, *104*(3), 347–363. https://doi.org/10.1111/j.2044-8295.2012.02127.x