Disentangling Syntactic and Lexical Complexity in Text Adaptation: Effects on Readers with Different Working Memory Capacity and Vocabulary Size

Sarubi Thillainathan (sarubi@lst.uni-saarland.de), Margarita Ryzhova (mryzhova@lst.uni-saarland.de), Alexander Koller, Vera Demberg
Department of Language Science and Technology, Saarland University

Text adaptation – the process of rewriting a text to tailor it to the specific needs of individual readers while preserving its meaning, for example through text simplification – typically modifies multiple linguistic dimensions simultaneously, including syntactic structure and lexical difficulty. As a result, existing studies provide limited insight into which exact aspects of adaptation drive changes in reading behavior and comprehension. To address this limitation, we constructed a controlled extension of the OneStopQA corpus [1] in which syntactic and lexical complexity are manipulated independently across four levels of syntactic complexity and four levels of lexical complexity.

We selected 18 advanced-level news texts from OneStopQA and generated adapted versions using a controllable text adaptation system from [4]. Syntactic rewrites manipulated dependency depth, whereas lexical rewrites manipulated lexical frequency. All automatically generated texts were subsequently refined manually to preserve meaning and ensure that the original comprehension questions remain answerable across all versions. Figures 1 and 2 show that the resulting texts successfully separate syntactic and lexical complexity dimensions, providing substantially greater experimental control than existing text adaptation datasets.

Using these materials, we present a planned experiment investigating how syntactic and lexical complexity affect reading behavior in adult L1 English readers, and whether these effects interact with individual differences in working memory capacity (WMC) and vocabulary size. Prior psycholinguistic research has linked WMC to the processing of long-distance dependencies and structurally demanding sentences [3], while vocabulary size has been associated with lexical processing efficiency and reduced sensitivity to word frequency effects [2]. However, these relationships have rarely been investigated using naturalistic text materials in which syntactic and lexical complexity are manipulated independently at the level of texts rather than isolated sentences, and across a continuous range of complexity levels rather than a binary original vs. simplified comparison.

In the planned study, participants will read the adapted texts in a self-paced reading paradigm and answer multiple-choice comprehension questions after each paragraph. Participants will additionally provide perceived difficulty ratings and complete standardized measures of WMC and vocabulary size. Reading times, comprehension accuracy, and perceived difficulty ratings will be analyzed using generalized linear mixed-effects models examining how different dimensions of linguistic complexity relate to reader-specific cognitive profiles. Data collection is currently ongoing, and the results will be presented at the conference.

References

[1] Yevgeni Berzak, Jonathan Malmaud, Omer Shubi, Yoav Meiri, Ella Lion, and Roger Levy. OneStop: A 360-participant English eye tracking dataset with different reading regimes. *Scientific Data*, 2025.

[2] Nina Mainz, Zeshu Shao, Marc Brysbaert, and Antje S Meyer. Vocabulary knowledge predicts lexical processing: Evidence from a group of participants with diverse educational backgrounds. *Frontiers in Psychology*, 8:1164, 2017.

[3] Bruno Nicenboim, Shravan Vasishth, Carolina Gattei, Mariano Sigman, and Reinhold Kliegl. Working memory differences in long-distance dependency resolution. *Frontiers in Psychology*, 6:312, 2015.

[4] Sarubi Thillainathan and Alexander Koller. Controllable text adaptation using in-context learning with linguistic features. In *AAAI2025 AI for Education - Tools, Opportunities, and Risks in the Generative AI Era*, 2025.

Figures and Tables

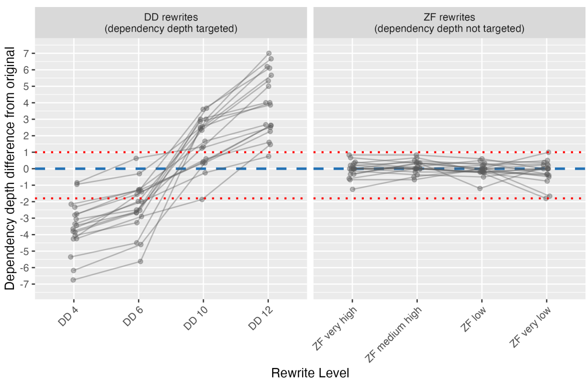

Figure 1: Dependency depth (DD) difference from the original advanced-level text across DD rewrites (left panel) and ZF (Zipf frequency) rewrites (right panel). Each line represents one story topic (N=18). In the left panel, x-axis labels indicate the targeted maximum dependency depth of sentences (DD 4, DD 6, DD 10, DD 12). In the right panel, x-axis labels indicate the target Zipf frequency of content words excluding named entities: very high = 8.0, medium high = 6.5, low = 3.0, very low = 1.0. Red dotted lines indicate the min/max DD change observed in ZF rewrites, where dependency depth was not targeted.

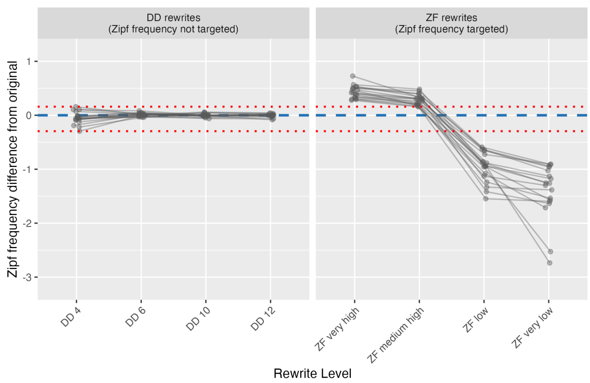

Figure 2: Zipf frequency (ZF) difference from the original advanced-level text across DD (dependency depth) rewrites (left panel) and ZF rewrites (right panel). Each line represents one story topic (N=18). In the left panel, x-axis labels indicate the targeted maximum dependency depth of sentences (DD 4, DD 6, DD 10, DD 12). In the right panel, x-axis labels indicate the target Zipf frequency of content words excluding named entities: very high = 8.0, medium high = 6.5, low = 3.0, very low = 1.0. Red dotted lines indicate the min/max ZF change observed in DD rewrites, where lexical frequency was not targeted.