Genre Effects on Eye Movements in Italian Reading: Evidence from the MultiplEYE_IT Dataset

Daniele P. Radicioni*[1], Patrizia Paggio[2], Alessandro Corona[2], Matteo Delsanto[1], Alberto Parola[2], Viviana Patti[1], Calogero Jerik Scozzaro[1]

[1]Computer Science Department, University of Turin – Italy
[2]Department of Nordic Studies and Linguistics (NorS), University of Copenhagen – Denmark

*contact: daniele.radicioni@unito.it

The MultiplEYE texts have been chosen to ensure coverage of different genres and text types (Kaspere *et al.*, 2026). We know, in fact, that eyetracking measures of reading vary depending on these features (Bolliger and Jäger, 2025). In this study we have analysed the four genres represented by the texts used to construct stimuli for the Italian collection to see whether they differ in style, lexical difficulty and syntactic complexity. For each text we have extracted type-token ratio (TTR), mean sentence length, mean word length and mean dependency distance (MDD). The measures show variability across the genres: the institutional texts display low lexical difficulty and high syntactic complexity; the popular scientific ones have the longest sentences and show great TTR variability; the literary texts have relatively high TTR and MDD; the argumentative ones score high on lexical and low on syntactic difficulty. We expect these differences to correspond to differences in the eyetracking measures. To test our expectations, we have analysed the data collected in Turin to extract the distribution of eyetracking measures reflecting early as well as later and cumulative processing. This analysis shows noticeable differences across the four genres. In particular, the trends we observe seem to indicate that genre has a limited influence on the earliest stage of processing (FPRT). By contrast, clearer genre-related variation emerges in the later and more cumulative measures (RRT, RPD_inc and TFT). At the same time, the sizeable variation in all the measures make it necessary to move beyond descriptive summaries and assess these apparent differences through inferential modelling that accounts for subject- and item-level variation. The analysis of genre effects indicates that reading times vary systematically across genres, but that these differences are not equally visible in all eye-tracking measures. Using mixed-effects regression models with random intercepts for subjects and items, we tested whether genre explains variation in reading behavior beyond baseline heterogeneity associated with individual readers and textual items. Literary texts were treated as the reference category, and separate models were fitted for first-pass reading time (FPRT), re-reading time (RRT), inclusive regression-path duration (RPD_inc), and total fixation time (TFT). The results show that genre effects are weakest in the earliest stage of processing and become more pronounced in later, more integrative measures. These findings suggest that genre affects not so much immediate word-by-word uptake as the amount of later rereading, revision, and cumulative effort required for interpretation.

Our results are consistent with recent literature showing that genre shapes eye-movement behavior at more global and strategic levels. Markevich and Streltsova (2025) report genre-driven differences in scanpath organization across narrative, expository, and poetic texts, highlighting the role of genre in shaping reading strategies beyond local fixation behavior. Our results are also compatible with the work by Bolliger and Jäger (2025), who show that text type modulates reading behavior and interacts with lexical predictors such as surprisal, word length, and frequency. Taken together, these studies converge with our findings in showing that genre is a determinant of reading behavior, while our results further suggest that its clearest effects emerge in rereading- and integration-sensitive measures rather than in early first-pass processing.

References

Bolliger L. S., Jäger L. A. (2025). Genre matters: How text types interact with decoding strategies and lexical predictors in shaping reading behavior. In *Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing*, Suzhou, China, pp. 7459–7476.

Kaspere R., Bondar A., ... (2026). The MultiplEYE text corpus: Towards a diverse and ever-expanding multilingual text corpus. LREC 2026: Language Resources and Evaluation Conference.

Markevich M., Streltsova A. (2025). The influence of text genre on eye movement patterns during reading. *Journal of Eye Movement Research 18*(6), 60.