# Old or Modern? A Computational Model for Classifying Poem Comprehension using Microsaccades

Patrizia Lenhart
patrizia.lenhart@student.uni-tuebingen.de
University of Tübingen
Tübingen, Germany

Enkeleda Thaqi
enkeleda.thaqi@tum.de
Technical University of Munich
Munich, Germany

Enkelejda Kasneci
enkelejda.kasneci@tum.de
Technical University of Munich
Munich, Germany

Nora Castner
nora.castner@uni-tuebingen.de
Human-Computer Interaction, University of Tübingen
Tübingen, Germany
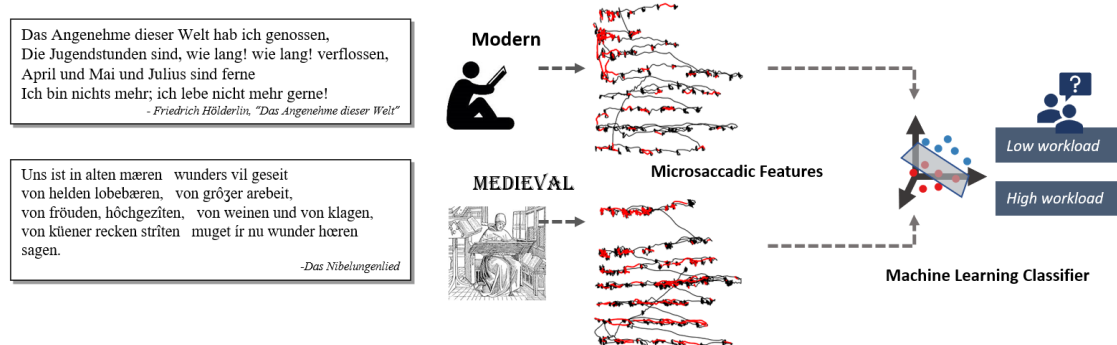
Figure 1: We looked at reading comprehension between poems written in modern high German and in Mittelhochdeutsch. We present an exploratory analysis of the feasibility of microsaccadic features for machine learning classifiers for reading comprehension. Public domain image featured: *Scriptorium Monk at Work* from Blades, William: Pentateuch of Printing with a Chapter on Judges (1891), https://www.fromoldbooks.org/Blades-Pentateuch/pages/scriptorium-monk-at-work/.

CCS CONCEPTS

• **Human-centered computing → Human computer interaction (HCI)**; • **Computing methodologies → Machine learning**.

KEYWORDS

Eye Movements, Microsaccades, Reading comprehension, Machine Learning

1 INTRODUCTION

The study of subtle eye movements such as microsaccades and pupil diameter gained a lot of scientific attention. So far, these events have been studied in the context of assessing the state of cognitive load performing different tasks. However, microsaccades to classify reading poetry in Modern and Middle High German (Mittelhochdeutsch) has yet to be researched. For many readers, Mittelhochdeutsch can appear to be a completely different language. Consequently, classifying workload while reading these poems could result in models with high prediction power. Figure 1 shows the workflow of this model.

2 RELATED WORK

Appel et al. and Krueger et al. have shown that eye movements such as change of pupil diameter and microsaccades serve as reliable indicators of cognitive workload during task performance. These studies found that the pupil diameter widens as the cognitive load increases, as well as, that the microsaccade rate decreases and the microsaccade magnitude increases. Krejtz et al. found that the microsaccade magnitude and pupillary response are sensitive to task difficulty. Yet, neither the microsaccade rate nor the relationship between microsaccade peak velocity and size were sensitive to task difficulty or time-on-task. Bowers and Poletti and Rima and Schmid

observed in their reading behavior studies that microsaccades are regressive movements intended to adjust for saccadic mislandings.

3 METHODS

We recorded eye movements using the Tobii Pro Spectrum at a sampling rate of 1200 Hz. Participants sat approximately 60 cm from the monitor and used a chin rest to minimize head movements. A 9-point calibration with 4-point validation was performed prior to poem reading.

The experimental session consisted of reading ten poems: Five modern and five Mittelhochdeutsch, which were presented in a randomized order. Following each poem, a comprehension question was presented as a motivation to actively process and engage with the poems.

A total of 47 participants, consisting of students from various semesters and study courses, were recruited for the study.

Data preprocessing involved removing datasets if the tracking ratio was below 80%. After this criterion, a total of 375 datasets were included in the final analysis. We smoothed the pupil diameter. Microsaccades were detected using an implementation of Engbert and Mergenthaler from the Engbert Microsaccade Toolbox [1]. We set a 6ms minimum duration and VFAC to 3 and added an additional parameter of maximum duration (30ms).

4 MODEL AND RESULTS

For the model and its evaluation, we worked on the assumption that the Mittelhochdeutsche poems would trigger higher mental workload. In line with previous research, we looked at microsaccade magnitude [Krejtz et al. 2018; Krueger et al. 2019]. However, we found microsaccade rate lead to overfitting for all models, thus we only report the magnitude. We used a 80-20 train-test split. The best performing models are found in Table 1.

Table 1: Test accuracy on different classification models

Model	Accuracy [%]
SVM	56.1
kNN	58.4
Decision Tree	**62.1**
Neuronal Network	58.1

As pupil dilation is also an indicator of workload [Appel et al. 2018; Just and Carpenter 1992], we tested for a significant difference in pupil diameter and found no significant differences ($p = 0.49$). Summary statistics also showed no significant differences for microsaccade rate ($p = 0.29$) and magnitude ($p = 0.9$).

We checked how engages the participants were during the task by assessing their responses. They answered correctly at a rate of 86.6% for modern poems and 60.9% for the Mittelhochdeutsche poems. We had expected higher scores for modern poem comprehension, thus it is unclear how engenged participants were.

5 CONCLUSION AND FUTURE OUTLOOK

We investigated the feasibility of microsaccade magnitudes as feature input for machine and deep learning classifiers to distinguish mental workload when reading poetry in both the Medieval German

language and Modern High German language. For Native Germans, Medieval German appears to be a whole different language and, therefore, we hypothesized that readers would have higher workload when reading Medieval German poetry. However, sadly this was not the case. First, we found no significant differences with regard to pupil dilation and reading both poem types indicating no differences in workload. Similarly, Menninghaus and Wallot found no differences in pupil dilation during poetry reading, though they were using it as an indicator for aesthetic appreciation.

Then, we further focused our investigation on machine learning approaches in an effort to extract potentially relevant information from the temporal behavior of the poetry readers. However, these models only slightly distinguished reading Modern from Medieval poems above chance level (62%). We chose microsaccade behavior as input for these models as previous literature has shown its link to workload [Krueger et al. 2019] and reading behavior [Bowers and Poletti 2017]. More importantly, microsaccades have not yet been investigated in the context of poetry reading, nor its interpretation.

From the current investigation, we can conclude that there were no differences in workload – as measured by microsaccades – when reading Medieval or Modern poetry. We suspect that our participants may not have put forth the effort towards fully comprehending the Medieval German poetry, which, we would like to point out, is a major limitation of the current study. Future iterations should better motivate people towards trying to interpret or even rate the poetry. As previous studies evaluating the aesthetics and reading "fluency" have found eye movement differences using summary statistics [Menninghaus and Wallot 2021; Xue et al. 2019], we still recommend microsaccade metrics as a new way to explore aspects of poetry interpretation. Moreover, We feel using data-driven, machine learning approaches to investigate how people read poetry can offer a greater extent of information not bound to poem specific details, rather common attributes of poem styles.

REFERENCES

Tobias Appel, Christian Scharinger, Peter Gerjets, and Enkelejda Kasneci. 2018. Cross-subject workload classification using pupil-related measures. In *Proceedings of the 2018 ACM Symposium on Eye Tracking Research & Applications*. 1–8.

Norick R. Bowers and Martina Poletti. 2017. Microsaccades during reading. *PLOS ONE* 12, 9 (09 2017), 1–14. https://doi.org/10.1371/journal.pone.0185180

Ralf Engbert and Konstantin Mergenthaler. 2006. Microsaccades are triggered by low retinal image slip. *Proceedings of the National Academy of Sciences* 103, 18 (2006), 7192–7197.

Marcel A Just and Patricia A Carpenter. 1992. A capacity theory of comprehension: individual differences in working memory. *Psychological review* 99, 1 (1992), 122.

Krzysztof Krejtz, Andrew T. Duchowski, Anna Niedzielska, Cezary Biele, and Izabela Krejtz. 2018. Eye tracking cognitive load using pupil diameter and microsaccades with fixed gaze. *PLOS ONE* 13, 9 (09 2018), 1–23. https://doi.org/10.1371/journal.pone.0203629

Eva Krueger, Andrea Chavaillaz, Rudolf Groner, Ben Sawyer, Andreas Sonderegger, and Peter Hancock. 2019. Microsaccades Distinguish Looking From Seeing. *Journal of Eye Movement Research* 12 (06 2019).

Winfried Menninghaus and Sebastian Wallot. 2021. What the eyes reveal about (reading) poetry. *Poetics* 85 (2021), 101526.

Samy Rima and Michael C. Schmid. 2021. Reading Specific Small Saccades Predict Individual Phonemic Awareness and Reading Speed. *Frontiers in Neuroscience* 15 (2021). https://doi.org/10.3389/fnins.2021.663242

Shuwei Xue, Jana Lüdtke, Teresa Sylvester, and Arthur M Jacobs. 2019. Reading shakespeare sonnets: combining quantitative narrative analysis and predictive modeling—an eye tracking study. *Journal of Eye Movement Research* 12, 5 (2019).

---

[1]https://github.com/lschwetlick/EngbertMicrosaccadeToolbox