

Summary

This thesis addresses the problem of detecting complex text by exploring whether recordings of readers' eye movements can be leveraged for learning what parts of texts obstruct readers, and investigates how this information can help improve NLP applications.

The problem of detecting and handling errors and deviations that make text unnecessarily difficult to read is becoming increasingly important to address, as the use of language technologies for improving information accessibility and communication efficiency grows.

In the thesis, four independent studies target the tasks of automatic text simplification, machine translation, sentence compression and lexical complexity detection. The empirical investigation presents evidence that it is possible to obtain and make use of information about text complexity from readers' gaze behaviour.

The results presented and discussed in the thesis contribute to the field of natural language processing by identifying important potentials and limitations of using gaze data in language technology.