Text Readability and Automatic Simplification
Connecting Linguistics, Processing and Educational Applications

Research Questions in the Thesis
1. Can we analyze the linguistic complexity of texts automatically?
2. Can we apply this analysis in Educational contexts?
3. Will text complexity affect a learner’s cognitive processing and performance outcomes?
4. Can we automatically simplify texts to a reading level appropriate to the learner?

Q1: Automatic Assessment of Text Complexity
- **Task**: Predicting the appropriate grade level for a text
- **Methods**: Texts: Texts written for learners at various grade levels
  1. Linguistic Variables studied:
     - Lexical: e.g., lexical density, type-token ratio
     - Syntactic: e.g., dependent clauses per sentence
     - Morphological: e.g., complexity of a word
     - Psycholinguistic: e.g., age of acquisition of words
  2. Modeling: machine learning (e.g., linear regression, support vector classification)
- **Results**: Our readability model is accurate, generalizable across texts and genres. (correlation: 0.9)

Q2: Application in Educational Contexts
1. **Reading Demands Project (German)**
   - **Aim**: Analyze the differences in linguistic complexity of German textbooks between schools, grades.
   - **Results**: (a) We can identify grade-wise and school-wise differences for certain linguistic features.
     (b) For predictive models, there are significant differences between publishers and schools.
     (c) Prediction is better at school level than grade level.
   - (funded by a LEAD Intramural Research Grant)

Q3: Impact of Text Complexity on Readers
- **Aim**: Understand how text complexity affects a learner’s cognitive processing and performance.
- **Methods**: Eye tracking, recall-comprehension questions and generalized additive mixed models.

Q4: Automatic Text Simplification
- **Aim**: perform text simplification as monolingual machine translation (using Moses toolkit).
- **Method**: use readability models for choosing difficult sentences to simplify. (Vajjala & Meurers, in prep.)
- **Example Simplification from my approach**: Hyper inflation is a condition where prices increase rapidly as a currency loses its value.
- **simplified**: Hyper inflation is a condition where prices increase very fast as money loses its value.

References