NLP System Demonstration

Xiaobin Chen

Tübingen University

May 17th, 2017
Flagstaff, AZ
A Web-based tool supporting automatic analysis of text.

- Four components:
  - corpus manager
  - feature selector
  - analysis generator
  - result visualizer

- Features:
  - Consistent, easy-to-use, friendly user interface
  - Modularized, reusable, and collaborative development of analysis components
  - Flexible corpus and feature management

- Use cases:
  - Complexity analysis / readability assessment
  - Authorship attribution
  - Plagiarism detection
  - ...
http://ctapweb.com/

Related publication:
Syntactic Benchmarks

Challenges learners in their individual Zone of Proximal Development using pedagogic developmental benchmarks of syntactic complexity.

- Modeling syntactic complexity development with a target language or pedagogic corpus (Newsela, 14,581 news articles in 9 developmental levels)
- Placement of user proficiency on the developmental benchmark
- Provision of comprehensible reading input, which is configurable in terms of the degree of challenge and the target grade level
- Supporting 14 syntactic complexity measures
Related Publication:
Form-Focused Linguistically Aware Information Retrieval

- Primary operations:
  - Web Search
  - Text Crawling
  - Parsing
  - Ranking

- Identifies the 87 grammatical constructions spelled out in the official English language curriculum of schools in Baden-Württemberg, Germany
System Demo—FLAIR

by Maria Chinkina & Madeeswaran Kannan supervised by Prof. Dr. Detmar Meurers
@ University of Tübingen, Germany | 2015-2017
Version 2.0

FLAIR is an online tool for language teachers and learners that:
- searches the web for a topic of interest
- analyzes the results for grammatical constructions and readability levels
- re-ranks the results according to your (pedagogical or learning) needs specified in the settings

http://samos.sfs.uni-tuebingen.de:8080/flair-2.0/

- Related publication:
NLP at Work?

- Which components/functions of these systems require NLP processing?
- What NLP processes are required to realize these functions?
- Think about your own teaching and research. How can the NLP technologies used in these systems be used to solve your problems?
Mean length of clause in tokens

Formula: \#tokens / \#clauses

- Sentence segmenter -> tokenizer -> count \#tokens
- Sentence segmenter -> tokenizer -> parser -> tree structure pattern matcher (Tregex) -> count \# of matches of the clause pattern -> \#clauses