

# Text Simplification

Tom WYSOCKI, Peter PRIBIL,  
Margot GUETTIER, Morgan RUIZ-HUIDOBRO

October 2021

## Presentation of the project

### Final goal:

Text A in, text A/B out.

# Current issues in text simplification

## Issues with the technology:

- Lack of data for training
- Rules are difficult to make
- Which rules to use?

## Notion in text comprehension

Structures which require higher brain activation:

- passive / active voices [1]
- multiple and double negation [2]
- relative clause (subject, object) [3]
- syntactic movement (Wh-fronting, inversion ...) [3]
- complex sentences
- PP?

# Complex Sentences

## Fewer words = happiness

- Higher brain activation on longer strings
- Poses difficulties on referencing

# Syntactic movement

## Linear ordering is good

- Transformed sentences proved to be more difficult
- High brain activation

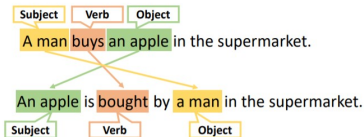
## PP Ambiguity

### Do we need it?

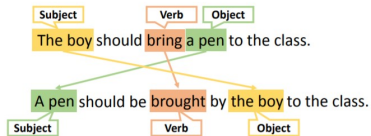
- It did not increase brain activation
- Bad argument: "exponentially increasing number of trees"

## Passive / Active voice sentences

Figure: Type examples of passive / active voice sentences[4]:



(a) The basic example



(b) The example of modal verb



## Multiple Negations

Figure: Different types of multiple negations sentences:

A. Single negation	The authors [that <b>the</b> critics recommended] have <b>never</b> received acknowledgment for a best-selling novel.
B. Multiple negation	The authors [that <b>no</b> critics recommended] have <b>never</b> received acknowledgment for a best-selling novel.
C. Double negation	<b>No</b> authors [that <b>the</b> critics recommended] have <b>never</b> received acknowledgment for a best-selling novel.

This type of sentences is quite rarely and we will treat it as the last priority.

## Existing systems

### **Syntactic simplification**

- MUSST
- Utilizing Prolog for converting between active sentence and passive sentence with three-steps conversion (by Trung Tran Quang)
- Splitting method used on complex Noun Phrases

### **Semantic simplification:**

- with BERT

### **System evaluation:**

- EASSE: Easier Automatic Sentence Simplification Evaluation

### **Other system:**

redalyc a web based TS system

## A different take on TS

### What is the purpose?

- Brain functionality vs intelligence?
- Simplification for the sake of simplicity?
- Shouldn't it be customizable?
- Time spent on reading?

## Timeline

**4 nov:**

Collection of syntactic simplification rules / coding

**22 nov:**

Cognitive linguistics research

**10 dec:**

Collection of semantic simplification rules / coding

**14 jan:**

Evaluation phase / coding

**15 fev:**

Defense preparation

- Presentation of the project
- Current issues in text simplification
  - Notion in text comprehension
    - Complexity
    - Movement
    - PP
  - Existing systems
    - Validity
    - Timeline
  - Questions

## Questions

Questions?



S. Feng, J. Legault, L. Yang, J. Zhu, K. Shao, and Y. Yang, "Differences in grammatical processing strategies for active and passive sentences: An fMRI study," *Journal of Neurolinguistics*, vol. 33, pp. 104–117, Feb. 2015.



P. Mukherjee, G. Leroy, D. Kauchak, S. Rajanarayanan, D. Y. Romero Diaz, N. P. Yuan, T. G. Pritchard, and S. Colina, "NegAIT: A new parser for medical text simplification using morphological, sentential and double negation," *Journal of Biomedical Informatics*, vol. 69, pp. 55–62, May 2017.



M. Poulsen and A. K. D. Gravgaard, "Who Did What to Whom? The Relationship Between Syntactic Aspects of Sentence Comprehension and Text Comprehension," *Scientific Studies of Reading*, vol. 20, pp. 325–338, July 2016.



E. Shetreet, N. Friedmann, and U. Hadar, “The Neural Correlates of Linguistic Distinctions: Unaccusative and Unergative Verbs,” *Journal of Cognitive Neuroscience*, vol. 22, pp. 2306–2315, Oct. 2010.