

# Text Simplification of College Admissions Instructions: A Professionally Simplified and Verified Corpus

Zachary W. Taylor, University of Southern Mississippi  
Maximus H. Chu and Junyi Jessy Li, University of Texas at Austin



## Why?

- L2 students have not accessed U.S. higher education at the same level as their English-fluent (native or L1) peers
- Kanno and Cromley (2013): using National Education Longitudinal Study (NELS: 1988)
  - 20% of L2 students were high school dropouts; 6% of L1 students dropped out
  - 12.5% of L2 students earned a bachelor's degree; 33% of L1 students
- Admission instructions for US higher education is at the level of **13-14th grades**

You must complete the following steps before USF will consider your application complete and begin admission evaluation. Pay the non-refundable \$30 application fee or submit an application fee waiver.

You must submit an online application with a nonrefundable \$30 application fee. You can also submit an application fee waiver.

## Simplification strategies

- Reducing syntactic complexity
  - Conjunctions, passive voice, etc
- Increasing lexical cohesion
- Elaborations and jargons
- Domain-specific principles
  - Removal of irrelevant sentences
  - Use of modal verbs & pronouns

## Expert Verification

- 10 subject-matter experts
- acceptability judgements = Accept or reject simplification
- All texts reviewed by three different SMEs—triangulate acceptability.

All conditionally accepted applicants must consent to, submit to and successfully complete a criminal background check through Certiphi Screening, Inc. Failure to do so will constitute failure to meet the pre-matriculation requirements established by SUNY Optometry and will result in the withdrawal of a conditionally accepted offer.

If you are conditionally accepted, you must consent to, submit to, and complete a criminal background check through Certiphi Screening, Inc. If you do not, we will withdraw your conditionally accepted offer.

- Text simplification: aims to help audiences read and understand a piece of text through lexical, syntactic, and discourse modifications, while remaining faithful to its central idea and meaning

## This Work: The PSAT Corpus

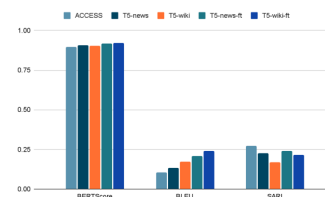
*Professionally Simplified Admissions Texts*



- A collection of admissions instructions from **112** higher education institutions in the US
- Professionally simplified:** first author (Dr. Zach Taylor), who has a doctoral degree in education and has worked professionally in US postsecondary education for over a decade, including work in undergraduate admissions.
- Professionally verified** by subject-matter experts

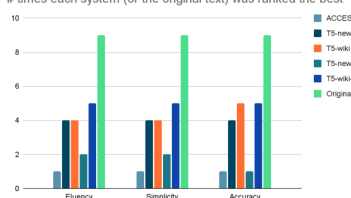
## Experiments

- Sentence-level simplification, using **1883 manual alignments**
- Models: ACCESS, T5 fine-tuned on News/Wiki
- Settings: Zero-shot vs. fine-tuning on PSAT



- Existing models do not achieve high BLEU/SARI
- Vanilla fine-tuned models a bit better at this domain

# times each system (or the original text) was ranked the best



- Experts prefer "not simplified" most of the time!
- Domain-specific principles not captured by models