# Russian Learner Translator Corpus 2.0
## https://rus-ltc.org

## RusLTC

**Parallel English<>Russian sentence-aligned corpus of multiple student translations.**

## Corpus size and metadata

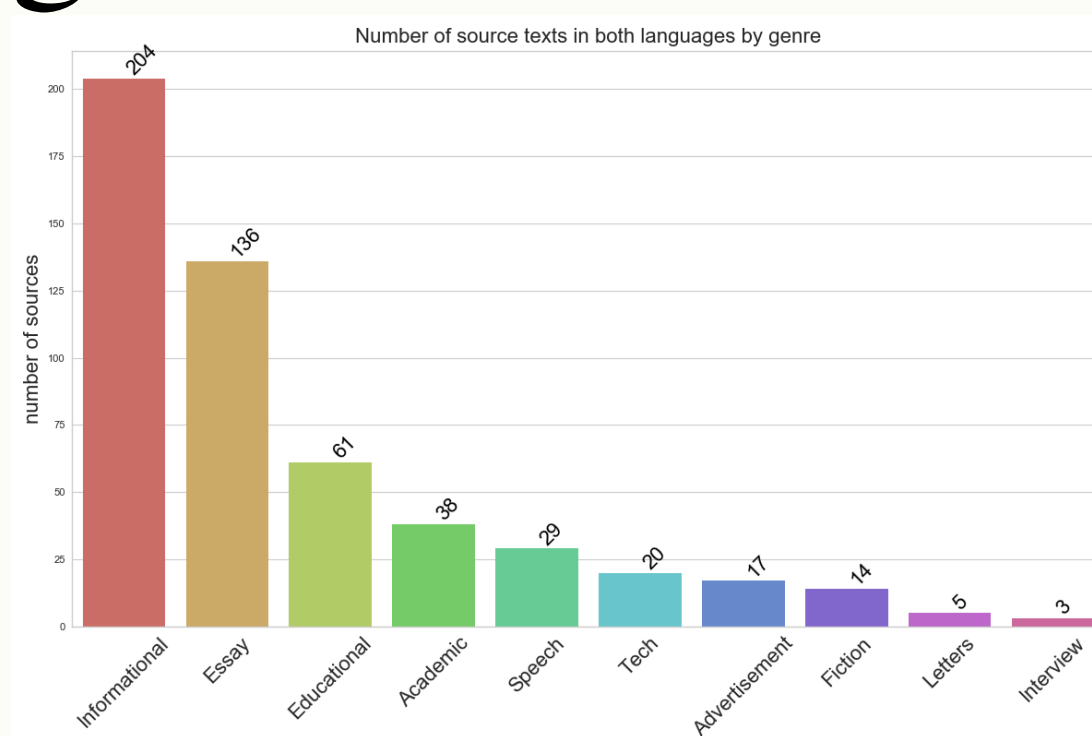► Basic statistics:
- 2.3 mln words, 4.8K texts;
- 402 English source texts, ≈ 8 Russian translations for each;
- RU targets: ≈ 380 words.

► Metadata:
- Translation grade;
- Draft / final translation (editing effort);
- Routine / Exam / Contest;
- At class / at home.

► Source text genres:



Number of source texts in both languages by genre

► Translators:
- Gender-annotated;
- 60% by advanced Translation Studies students;
- 14 Russian universities.

## Formats and availability

► Plain text documents;
► TMX (XML-based bitext format);
► stand-off error annotation files;
► fully public, downloadable, CC-BY-SA.

## Error-annotated translations

► > 750 texts, 300K words, 16K error labels
► 30 error types, 3 weights, 10 tech attributes

## New search interface: ElasticSearch + Tsakonian

## RusLTC at RANLP-2019

Research Group in Computational Linguisstics, UK