

Building a learner corpus for Russian

www.web-corpora.net/RLC

Russian Learner Corpus

- **open** learner corpus for Russian
- **search** over lemma, grammatical features, and error tags
- **clear distinction** between HL (heritage language) and L2 (second language)
- **multifunctional** platform that provides a single interface for uploading, annotating and search

www.web-corpora.net/RLC

Statistics

Texts	3 767
Sentences	58 483
Words	749 314

Annotated errors	23 809
------------------	--------

Language background

heritage	1 579
foreign	1 981
unknown	207

- 9 dominant languages
- partners from 12 countries
- partners from 20 universities

www.web-corpora.net/RLC/stats

Dominant languages

English	2754 (85%)	Korean	98
German	219	Kazakh	82
French	177	Finnish	64

Also:

Italian, Dutch, Norwegian, Japanese, Serbian

www.web-corpora.net/RLC/stats

User hierarchy

Staff	Add new users Mark texts as checked
Contributor	Upload new texts Add metadata
Annotator	Annotate texts online: <ul style="list-style-type: none">- add error tags- add corrections- add comments View full texts Mark texts as annotated
Guest (no registration)	Search by: <ul style="list-style-type: none">- lemma- PoS- grammatical features- error tags Select subcorpora

Building a learner corpus for Russian

Thank you!

www.web-corpora.net/RLC