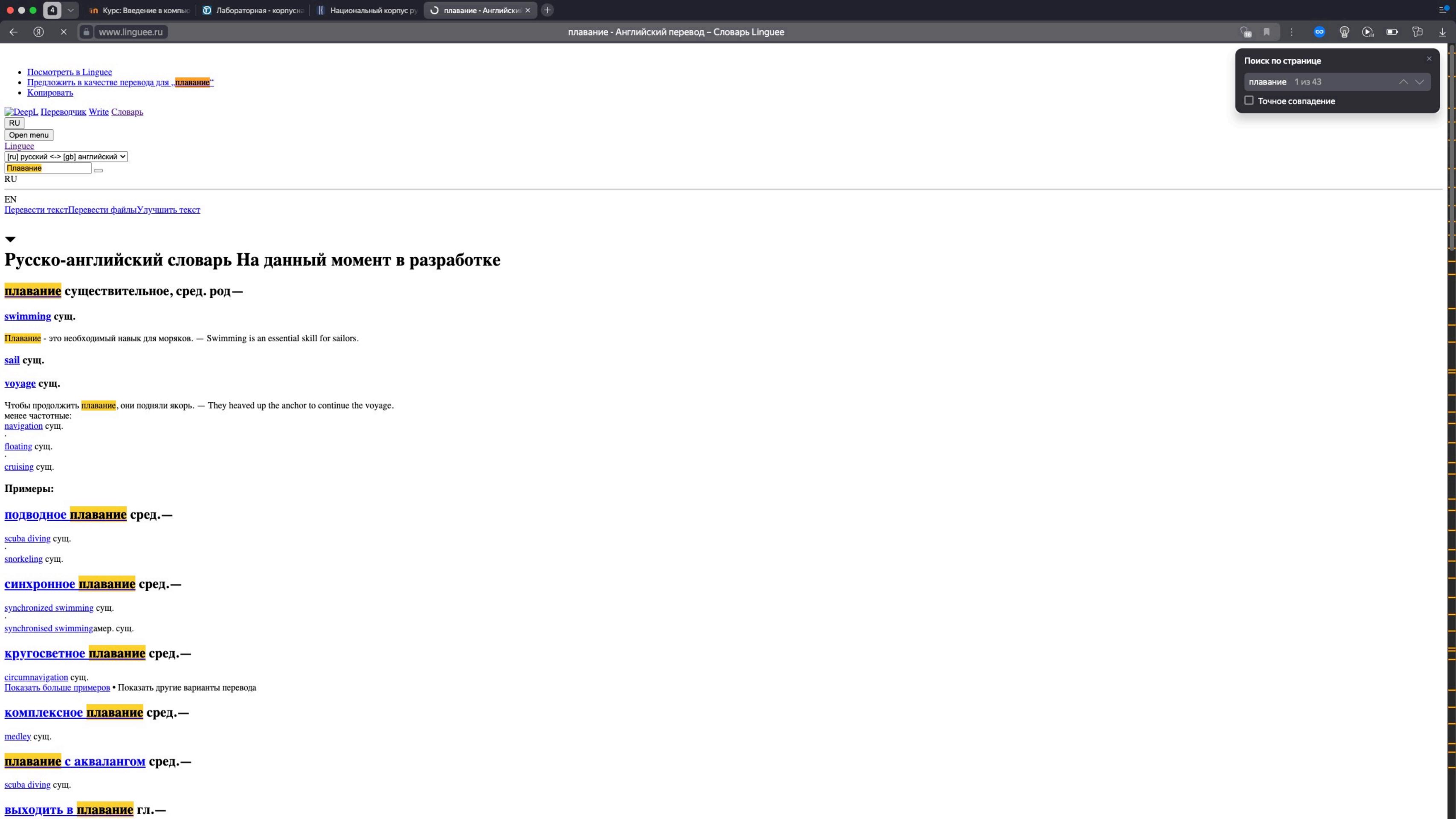
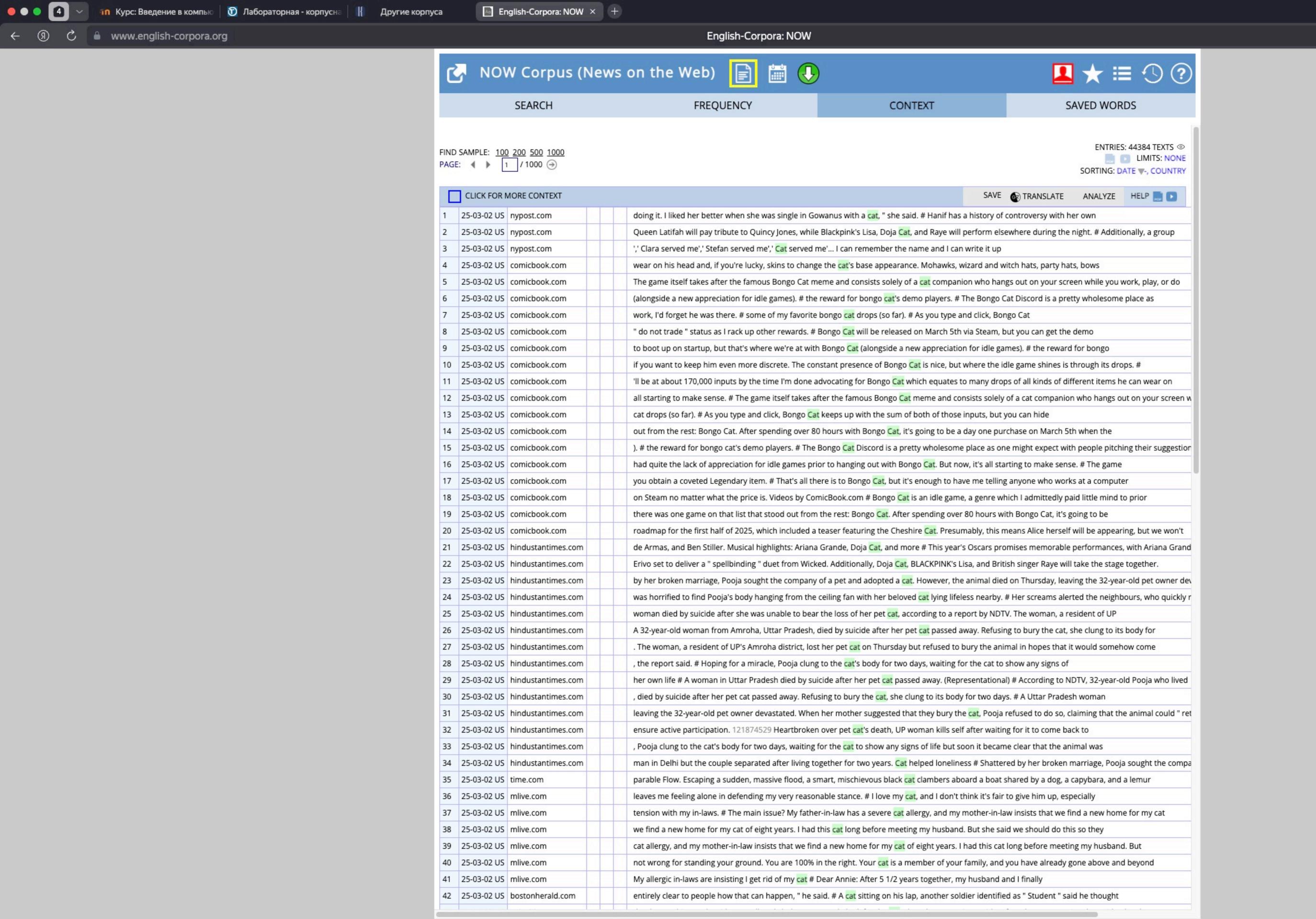
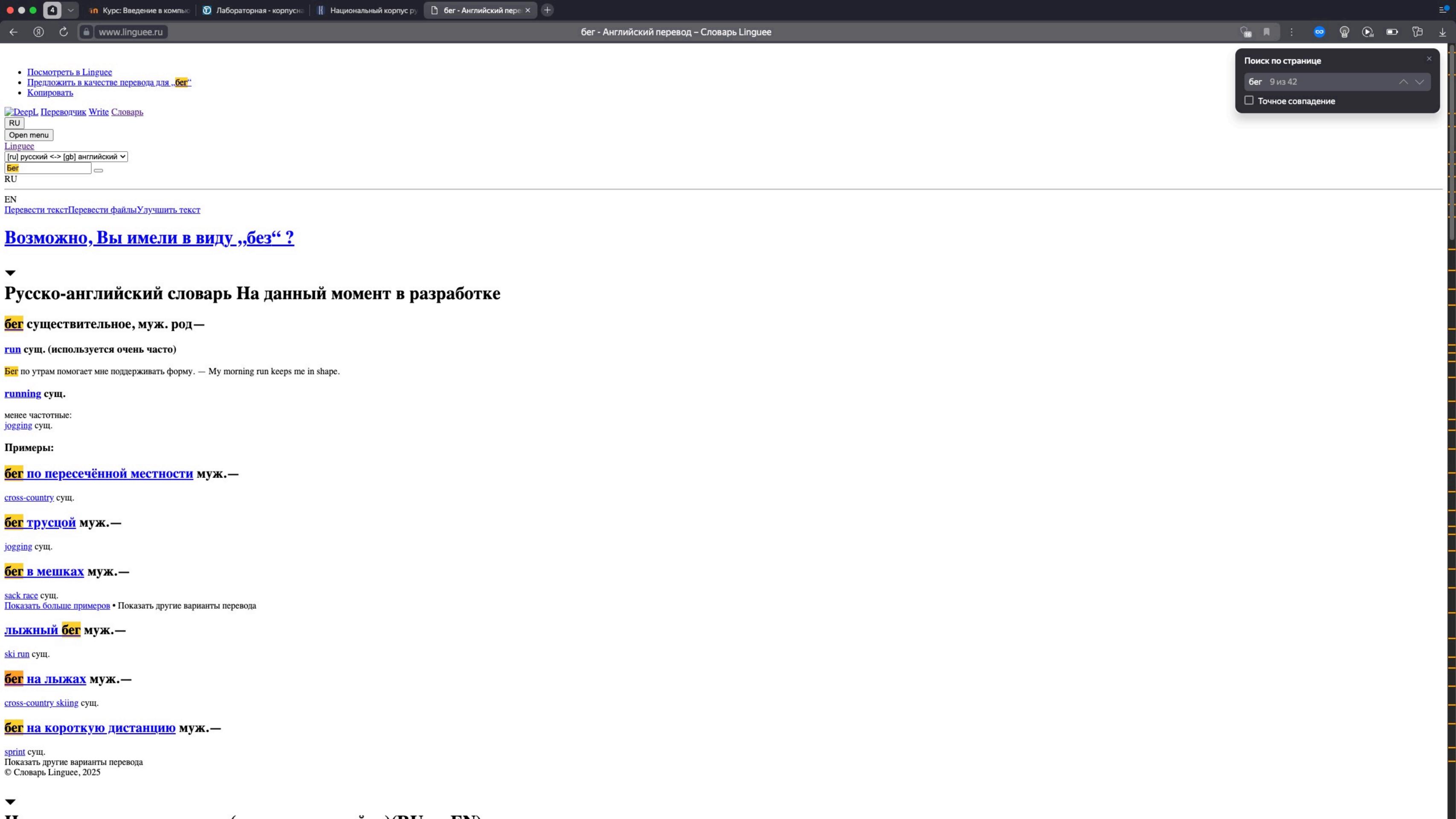


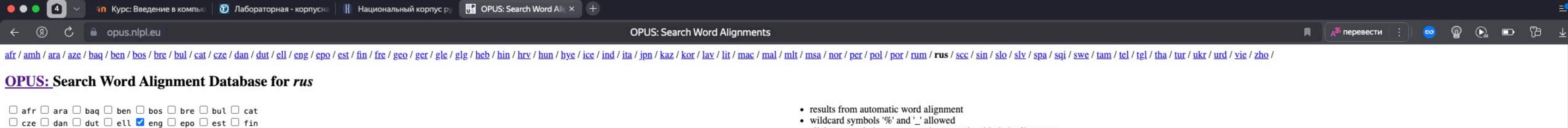
№ 125% :





<u>∞</u> @ [≥] □ [∃ √

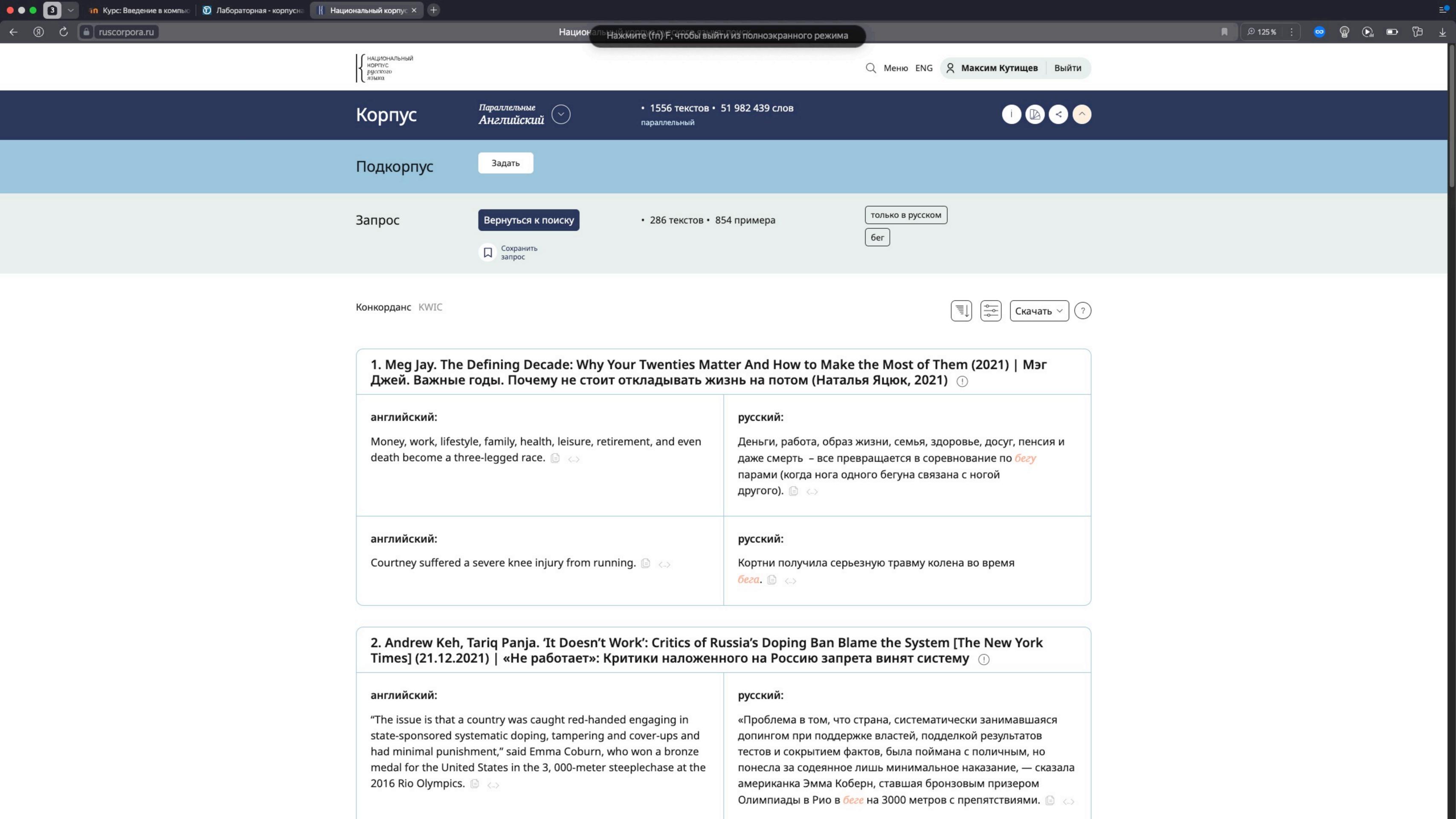




☐ fre ☐ geo ☐ ger ☐ glg ☐ heb ☐ hin ☐ hrv ☐ hun
□ hye □ ice □ ind □ ita □ jpn □ kaz □ kor □ lav
□ lit □ mac □ mal □ msa □ nor □ per □ pol □ por
□ rum □ scc □ sin □ slo □ slv □ spa □ sqi □ swe
□ tam □ tel □ tgl □ tha □ tur □ ukr □ urd □ vie
□ zho
Eer select from ● all ○ Books ○ DGT ○ Europarl ○ News-Commentary11 ○ OpenSubtitles2018 ○ SETIMES2 ○ TED2013 ○ Wikipedia

prob eng 52 (0.06) Running

- wildcard symbols '%' and '_' allowed
 click on translations to query these words with their alignments
 click on frequencies to get concordance lines from the corpus (max 100)
 the concordancer does not use word alignment





English-Corpora.org

corpora PDF guides videos related resources users my account upgrade help

Overview: brief | detailed D

Expanded help files

These are the most widely used online corpora, and they serve many different purposes for teachers and researchers at universities throughout the world. In addition, the corpus data (e.g. full-text, word frequency) has been employed by a wide range of companies in many different fields, especially technology and language learning.

The links below are for the free online interface. You can also purchase and download 🕦 the corpora for use on your own computer.

Corpus	Download	# words	Dialect	Time period	Genre(s)
News on the Web (NOW)	•	20.3 billion+	20 countries	2010-yesterday	Web: News
iWeb: The Intelligent Web-based Corpus	•	14 billion	6 countries	2017	Web
Global Web-Based English (GloWbE)	0	1.9 billion	20 countries	2012-13	Web (incl blogs)
Wikipedia Corpus	0	1.9 billion	(Various)	2014	Wikipedia
Coronavirus Corpus	0	1.5 billion	20 countries	2020-2023	Web: News
Corpus of Contemporary American English (COCA)	0	1.0 billion	American	1990-2019	Balanced
Corpus of Historical American English (COHA)	0	475 million	American	1820-2019	Balanced
The TV Corpus	0	325 million	6 countries	1950-2018	TV shows
The Movie Corpus	0	200 million	6 countries	1930-2018	Movies
Corpus of American Soap Operas	0	100 million	American	2001-2012	TV shows
Hansard Corpus		1.6 billion	British	1803-2005	Parliament
Early English Books Online		755 million	British	1470s-1690s	(Various)
Corpus of US Supreme Court Opinions		130 million	American	1790s-present	Legal opinions
TIME Magazine Corpus		100 million	American	1923-2006	Magazine
British National Corpus (BNC) *		100 million	British	1980s-1993	Balanced
Strathy Corpus (Canada)		50 million	Canadian	1920s-2000s	Balanced
CORE Corpus		50 million	6 countries	2014	Web
From Google Books n-grams (compare)					
American English		155 billion	American	1500s-2000s	(Various)
British English		34 billion	British	1500s-2000	(Various)

