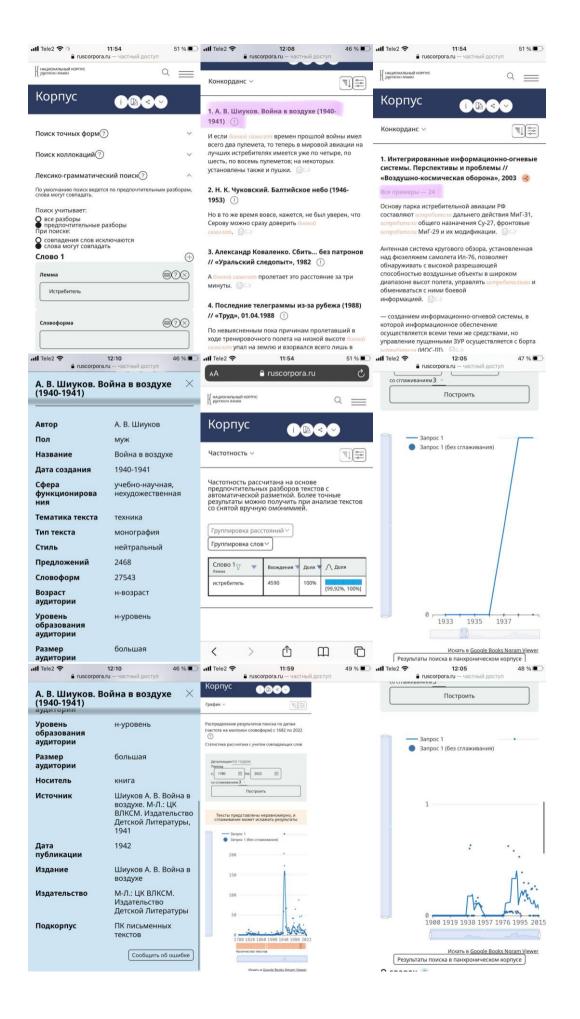
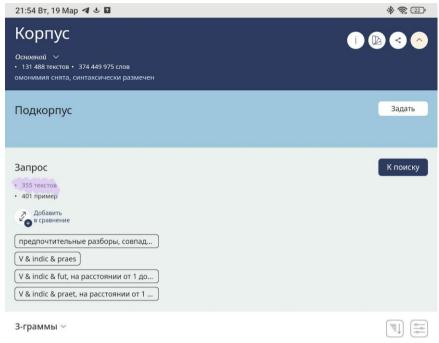
Уровень 3:

Задания по НКРЯ



2-20:

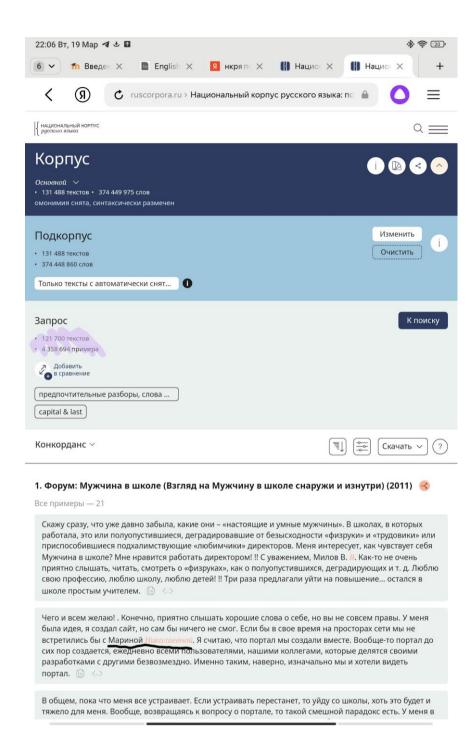
3.

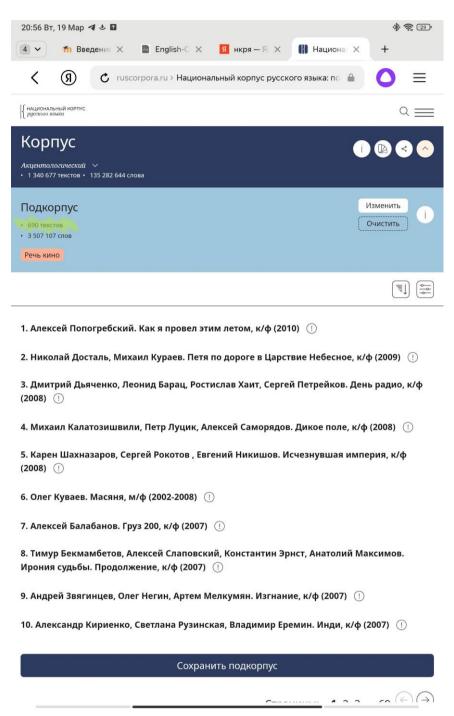


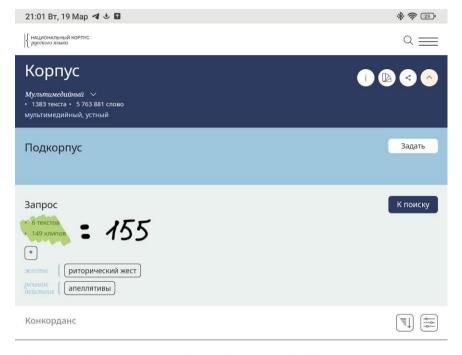
N-граммы рассчитаны на основе словоформ и не учитывают совпадающие слова

Nο	Тексты 🔻	Вхождения ▼	ipm ▼	Фрагмент	
1	3	4	0.01	может хватит спросил	
2	2	2	0.01	значит будет сказал	
3	2	2	0.01	может зайдем предложил	
4	2	2	0.01	может пригодится сказал	
5	2	2	0.01	может выпьем предложил	
6	1	1	0	бошиься исключим согласился	
7	1	1	0	бошиься уйдет говорила	
8	1	1	0	бошиься украдут спросил	
9	1	1	0	боюсь помрет жаловалась	
10	1	1	0	боюсь уйдет сказал	
11	1	1	0	боюсь украдут сказал	
12	1	1	0	бывает будет было	
13	1	1	0	верует спасется прибавил	
14	1	1	0	вижу опоздаем сказал	

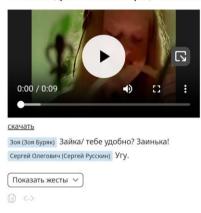
5.







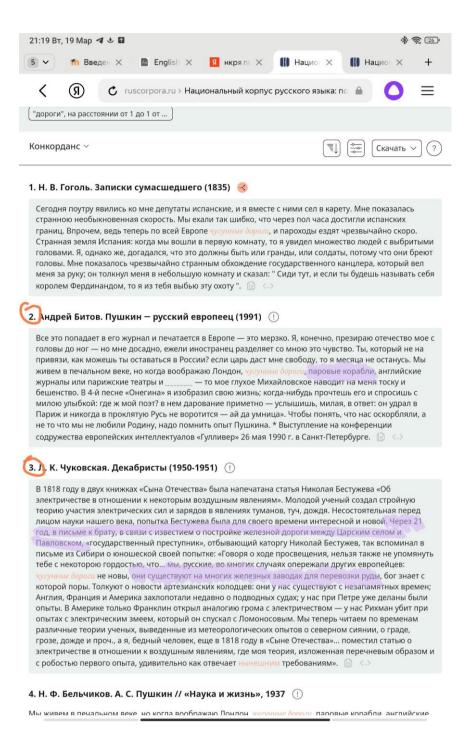
1. Александр Рогожкин. Операция С Новым Годом, к/ф (1996)



2. Александр Рогожкин. Операция С Новым Годом, к/ф (1996)



18.



20. Конструкция "интересоваться кем-либо" распространена намного больше, чем конструкция "интересоваться о ком-либо", поскольку примеры второго почти не были встречены

21.

Результаты поиска по запросу:	нкря, основной корпус			подкорп	ус худ. лите	ратуры XI)	Κв.		
	размер:	азмер: 283 431 966			размер: 29 796 584				
	кол-во	абсолют.		кол-во	абсолют.				
	текстов	частота	ipm	текстов	частота	ipm			
махает	92	1111	x	12	34	x			
машет	-7386	10	x	121	X	x			
	х = абсолютная частота/размер корпуса* 1 000 000								

Уровень 2:

Ответы на вопросы из 11 Лабораторная работа № 11 «Корпусная лингвистика. Корпус современного американского варианта английского языка (COCA)»

- СОСА создан профессором Марком Дэвидсом
- Корпус характеризуется большим объемом текстов(книг, статьи, рецензии, блоги, новостные статьи) за последние десятилетия, которые отражают современный английский язык
 - Разметки:
 - Лингвистическая выделение каждого слова тегом, указывающим его часть речи и другие лингвистические характеристики
 - Синтаксическая анализ структуры предложения и установка связи между словами, выявляя отношения между словами
 - Семантический выделение каждого слова(фразы) тегом, указывающим на отношение слов в контексте
 - Дискурсивная отражение связи между предложениями и абзацами в тексте
 - Морфологическая падеж, число, время
 - Леммантическая приведение слова к его лемме для упрощения анализа и поиска данного слова в тексте
 - 1. Поиск контекста: Это поможет им понять, как используются определенные слова или фразы в различных контекстах.
 - 2. Поиск синонимов и антонимов: полезно при поиске наилучшего перевода.
 - 3. **Изучение структуры предложений**: помогает переводчикам лучше понять грамматические конструкции и особенности синтаксиса в английском языке.
 - 4. **Изучение употребления идиом и фразеологизмов**: помогает переводчикам понять и правильно передать их значение на другой язык.
 - 5. Проверка специфической лексики и терминологии: СОСА может быть полезным для переводчиков, работающих с определенной тематикой или областью знаний, чтобы проверить использование специфической лексики и терминологии в контексте.

6. **Сравнение стилей и жанров**: позволяет переводчикам сравнивать различные стили письма и жанры текстов для выбора соответствующего перевода.

COCA (Corpus of Contemporary American English) включает в себя несколько видов подкорпусов, каждый из которых представляет собой набор текстов, собранных и разделенных по определенным критериям. Вот описание каждого из подкорпусов COCA:

- 1. **Spoken (Разговорный)**: состоит из устного английского языка(интервью, диалоги, радиопередачи и т.д.) Этот подкорпус отражает естественное использование языка в разговорной речи и помогает исследователям и переводчикам понять особенности устного английского.
 - 2. **Fiction (Художественная литература)**: включает тексты художественной литературы, такие как романы, рассказы, повести и т.д. Этот подкорпус помогает изучать структуру рассказов, использование литературных приемов и особенности стиля в художественных текстах.
 - 3. **Magazine (Журналы)**: содержит тексты из различных журналов, включая научные, популярные, модные и другие издания. Этот подкорпус предоставляет информацию о темах, стилях и лексике, используемых в журнальных статьях.
 - 4. **Newspaper (Газеты)**: включает тексты из различных газет и новостных изданий. Этот подкорпус позволяет исследователям анализировать новостные статьи, обсуждения актуальных событий и использование специфической журналистской лексики.
 - 5. **Academic (Научные статьи)**: содержит научные статьи и тексты из различных академических изданий. Этот подкорпус предоставляет информацию о специализированной терминологии, структуре научных текстов и особенностях академического стиля письма.