

Горайнова Анастасия,173-1

Отчет.

Нw5

Отчет по Заданию №1 проведу на примере слова «ныне»:

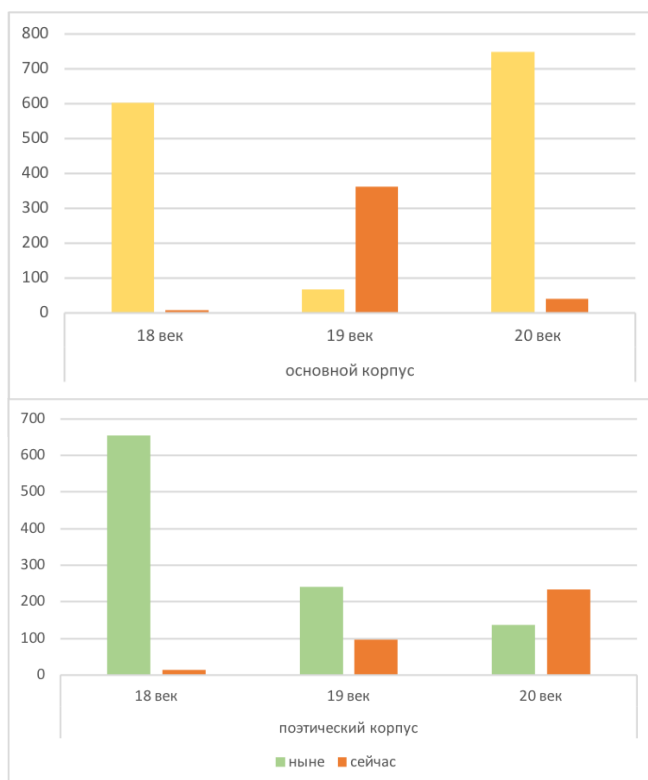
1. На сайте Национального корпуса русского языка я выбрала основной корпус.
2. Задала и сохранила подкорпус (пункт: год создания : с 1700 по 1799, отметив графу «точное вхождение»)
3. В поисковую графу «слово или фраза» вбила слово «ныне».
4. Было выдано 2 753 вхождения и 4 565 776 слов(объем подкорпуса).
5. Далее я высчитала частоту использования этого слова по формуле : $\frac{\text{кол-во вхождений}}{\text{объем подкорпуса}} * 1000000$.
6. Полученное число я записала в ячейку таблицы.
7. Такой же порядок действий был проделан с 19,20 веками в основном и поэтическом корпусе и со словом «сейчас».
8. Закончив заполнение таблицы, я выделив по отдельности каждую таблицу, создала две гистограммы (кликнув на значок «вставка»).

Выводы:

Основной корпус: слово «сейчас» употреблялось чаще всего в 20 веке, а «ныне» в 18. В 19 веке можно наблюдать увеличение популярности слова «сейчас», в то время, как словом «ныне» пренебрегают. Слово «ныне» было менее популярно, чем слово «сейчас» в 19 и в 20 веках.

Поэтический корпус: слово «ныне» употреблялось примерно в равном количестве в 19 и в 20 веках, а в 18 веке использовалось с наибольшей частотой. Говоря о слове «сейчас», невозможно сказать, что оно употреблялось одинаково в эти временные промежутки(18,19,20вв.). Гистограмма показывает увеличение частоты использования начиная от единичных случаев в 18 веке и заканчивая уже популярной тенденцией использования в 20 веке.

	основной корпус		
	18 век	19 век	20 век
ныне	602,9643154	66,86918367	748,478854
сейчас	8,103770312	362,1947064	38,9019048
	поэтический корпус		
	18 век	19 век	20 век
ныне	654,237258	241,28766	137,440845
сейчас	13,194701	97,1381074	234,169814



Задание №2:

1. Для выполнения этого задания необходимо работать в основном корпусе.
2. Я задала подкорпус, указав период правление Екатерины Второй(1762-1796гг.) и сохранила его.
3. Я использовала «лексико-грамматический поиск», вбивая по 1 слову в строку «слово»
4. Я стала вбивать определенные слова в И.п. , указывая дополнительные признаки в графе семантических/грамматических признаках.

Работа с первым диалогом:

Я решила вбить слово «поганка» в поисковую графу, так как мне показалось, что именно по этому слово можно будет понять, относиться ли первый отрывок к периоду правление Екатерины Второй.

Я была права: по моему запросу не было найдено ни одного документа или вхождения с упоминанием этого слова. Хотя поиск велся по подкорпусу объемом 949 документов, 173 266 предложений, 3 271 533 слова.

На этом основании я сделала вывод, что первый отрывок не мог встретиться в таком виде, в каком был представлен в задании в период правления Екатерины Второй.

Работа со второй репликой:

Снова создала подкорпус, с тем же периодом-правлением Екатерины Второй.

И решила искать слово «лопух», указать все необходимые грамматические признаки, например творительный падеж.

Результат: по этому запросу ничего не найдено. Хотя поиск велся по подкорпусу объемом 949 документов, 173 266 предложений, 3 271 533 слова.

На этом основании я сделала вывод, что и вторая реплика не могла встретиться в таком виде, в каком был представлен в задании в период правления Екатерины Второй.

Работа с третьей репликой:

Решила начать поиск со слова «растолковать», указав в грамматических признаках (будущее время, первое лицо и др). Было найдено 4 документа, 5 вхождений. Поиск велся по подкорпусу объемом 949 документов, 173 266 предложений, 3 271 533 слова.

Пример:

А. П. Сумароков. Приданое обманом (1769)

«Я, сударь, вас, ежели вы прикажете, от страха смертного совсем освобожу и о болезнях вашей лучше всех докторов **растолкую**.»

Далее я решила вбить в поиск слово «напасть», также указав необходимые грамматические признаки. Было найдено 46 документов, 70 вхождения, поиск велся по такому же подкорпусу с таким же объемом.

Пример:

Д. И. Фонвизин. Опыт российского сословника (1783-1784)

«Дураку закон не писан. Несчастье, **напасть**, беда, бедствие Все сии слова возвещают и знаменуют злключения; но несчастье называется всякое злое происшествие; **напасть** же есть злключение нечаянное.»

Задание №3(бонусное):

1. Я выбрала основном корпус и создала подкорпус, указав дату создания 1950 год и вплоть до настоящего времени, сохранила подкорпус.
2. В поиск вбила слово «щас».
3. Я открыла результаты, перешла к графику, а также открыла таблицы, скопировала таблицу «всего», скопировала ее.(распределение по годам(частота на миллион словоформ).
4. Вставила в эксель, отсортировала по убыванию, ориентируясь на столбец «использовано».
5. Я решила построить график по первым 10 годам таблицы.

К сожалению, у меня не получилось его построить, но я решила все же вставить свою отсортированную таблицу, по ней можно понять и увидеть пик популярности использования слова «щас».

То же самое я проделала с доп.признаками жанра(три:детектив,боевик;фантастика;юмор и сатира). Точно также я сортировала таблицы частоты использования слова в каждом жанре, выбирала год, в котором частота наибольшая и создала график.

	Всего
год	использовано
1984	1083377.589386
2002	10908038.230927
1986	1110933.883412
1964	1127323.725665
1980	1148249.157224
1972	1153921.682647
1990	1243136.640596
1978	1299567.784321
1982	1302433.542849
2014	1316853.571429

Пик использования слова "щас" в определенных жанрах текста

