

Домашняя работа №5.

Задание 1.

Частоты употреблений слов «ныне» и «сейчас» в 18, 19 и 20 в. можно вычислить по формуле: количество вхождений/количество слов необходимого корпуса, заданного на промежуток каждого века*миллион, т.к. при распределении по годам частота рассчитывается на миллион словоформ.

Частота употребления слов «ныне» и «сейчас» в основном корпусе:

18 век

Ныне: $2\,914 / 5\,202\,929 * 1\,000\,000 = 560$

Сейчас: $43 / 5\,202\,929 * 1\,000\,000 = 8,2$

19 век

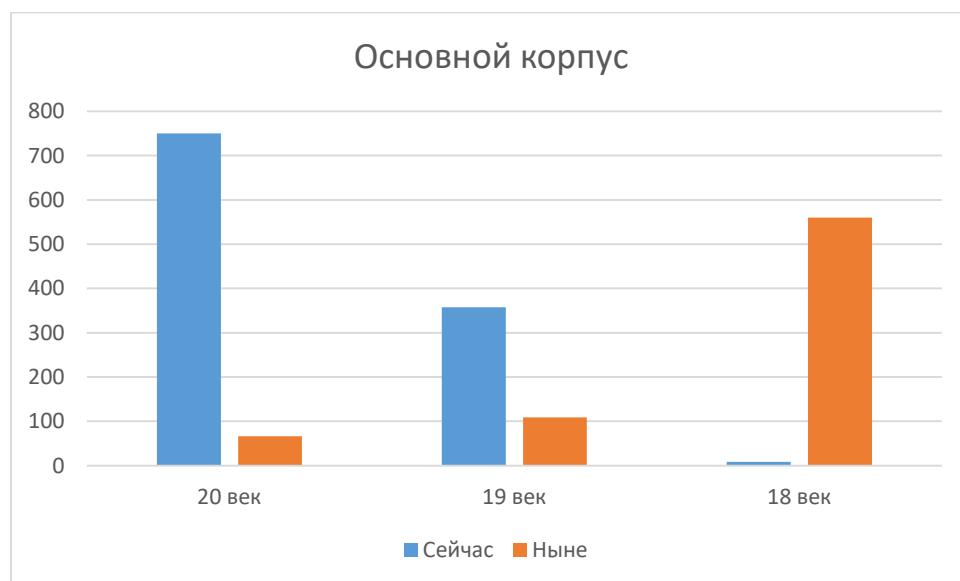
Ныне: $5\,922 / 54\,274\,217 * 1\,000\,000 = 109,1$

Сейчас: $19\,409 / 54\,274\,217 * 1\,000\,000 = 357,6$

20 век

Ныне: $10\,071 / 152\,456\,773 * 1\,000\,000 = 66$

Сейчас: $114\,368 / 152\,456\,773 * 1\,000\,000 = 750,1$



Вывод: от 18 к 20 веку прослеживается тенденция к учащению употребления слова «сейчас» и, наоборот, снижение частоты употребления слова «ныне».

Частота употребления слов «ныне» и «сейчас» в поэтический корпус:

18 век

Ныне: $595 / 909\,456 * 1\,000\,000 = 555,2$

Сейчас: $12 / 909\,456 * 1\,000\,000 = 13,1$

19 век

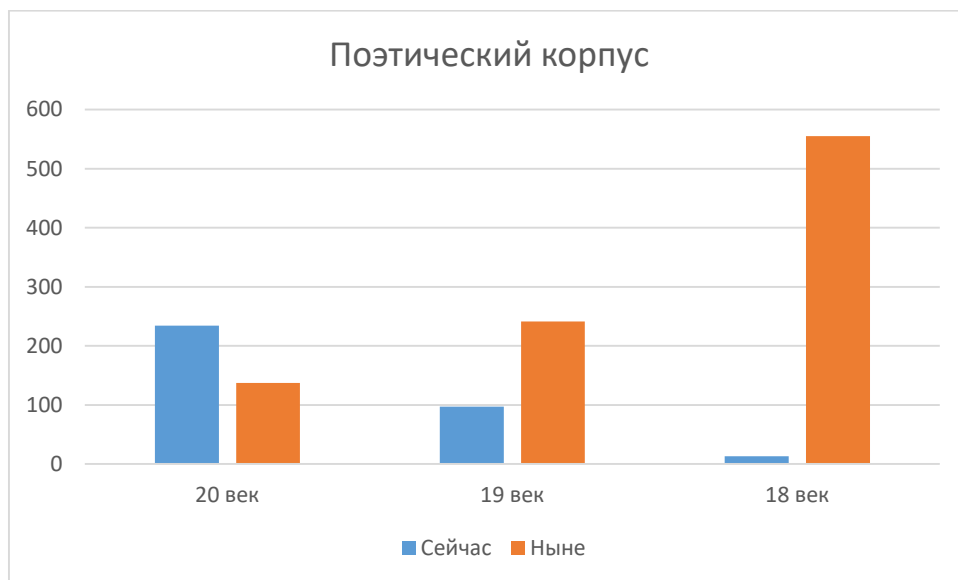
Ныне: $852/3\ 531\ 055 \cdot 1\ 000\ 000 = 241,2$

Сейчас: $343/3\ 531\ 055 \cdot 1\ 000\ 000 = 97,1$

20 век

Ныне: $898/6\ 533\ 720 \cdot 1\ 000\ 000 = 137,4$

Сейчас: $1\ 530/6\ 533\ 720 \cdot 1\ 000\ 000 = 234,1$



Вывод: в поэтическом корпусе прослеживается та же тенденция к учащению употребления слова «сейчас» и снижению частоты употребления слова «ныне» от 18 к 20 веку, но, в то же время, в поэтическом корпусе эта тенденция к 20-му веку становится не такой сильно выраженной, в отличие от ее проявления в основном корпусе: в 19 и 20 веке в поэтическом корпусе не существует такого большого разрыва в частотах употребления слов «ныне» и «сейчас».

Задание 2.

Времена правления Екатерины II приходятся на период с 1762 по 1796.

Чтобы определить, могли ли эти диалоги встретиться во время правления императрицы, следует задать временной подкорпус создания текста с 1762 по 1796 год и найти слова, особенно выделяющиеся своей формой.

1-ый диалог: по корпусу мы находим 207 употреблений словосочетания «Александр Невский», 1 употребление слова «пряжка». Не найдено документов со словами «поганка», «сволочь», словосочетанием «пиши пропало». Орден Александра Невского существовал с 18 по 20 век, так что большое количество вхождений словосочетаний «Александр Невский» не может являться ключевым показателем языка времени правления императрицы. Из этого следует, что существование этого диалога в период правления Екатерины II крайне маловероятно.

2-ой диалог: находим 49 вхождений междометий «ох», но ни одного вхождения слова «лопух», являющегося ключевым для этой фразы. Следовательно, этот диалог тоже, скорее всего, не мог существовать в период правления императрицы.

3-ий диалог: находим 12 вхождений слова «растолковать», 133 употребления слова «гнев», 44 употребления слова «напасть». Ключевые слова диалога, как и сам диалог теоретически могли встретиться во время правления Екатерины II, хотя мы и знаем, что сама фраза принадлежит произведению 1824 года публикации.