**Отчёт по проделанной работе по теме: «Использование параллельного корпуса для количественного изучения лингвоспецифичной лексики».**

Для выполнения данного задания мне требовалось выбрать два слова: одно лингвоспецифичное слово, а другое неспецифичное. Я выбрала два слова: «злобный» и «дыня». Прежде чем проанализировать их через сайт НКРЯ, я выдвинула гипотезу, что слово «злобный» является лингвоспецифичным, так как скорее всего его можно перевести огромным количеством слов и выражений на английский язык, в отличие от слова «дыня». Национальный корпус русского языка и формулы в самом задании помогли в подтверждении моей гипотезы.

**1)** Начну со слова «злобный». После того, как я вбила его на сайте, мне открылась страница, из которой видно, что: слово имеет 72 вхождения и огромное количество различных примеров. Я провела подсчёт, и оказалось, что данное слово имеет 37 различных моделей перевода, самая частотная повторяется 12 раз («malicious»), а вторая по частоте – 11 раз («evil»). Уже можно заметить, что слово имеет огромное количество моделей, что является одной из черт лингвоспецифичного слова.



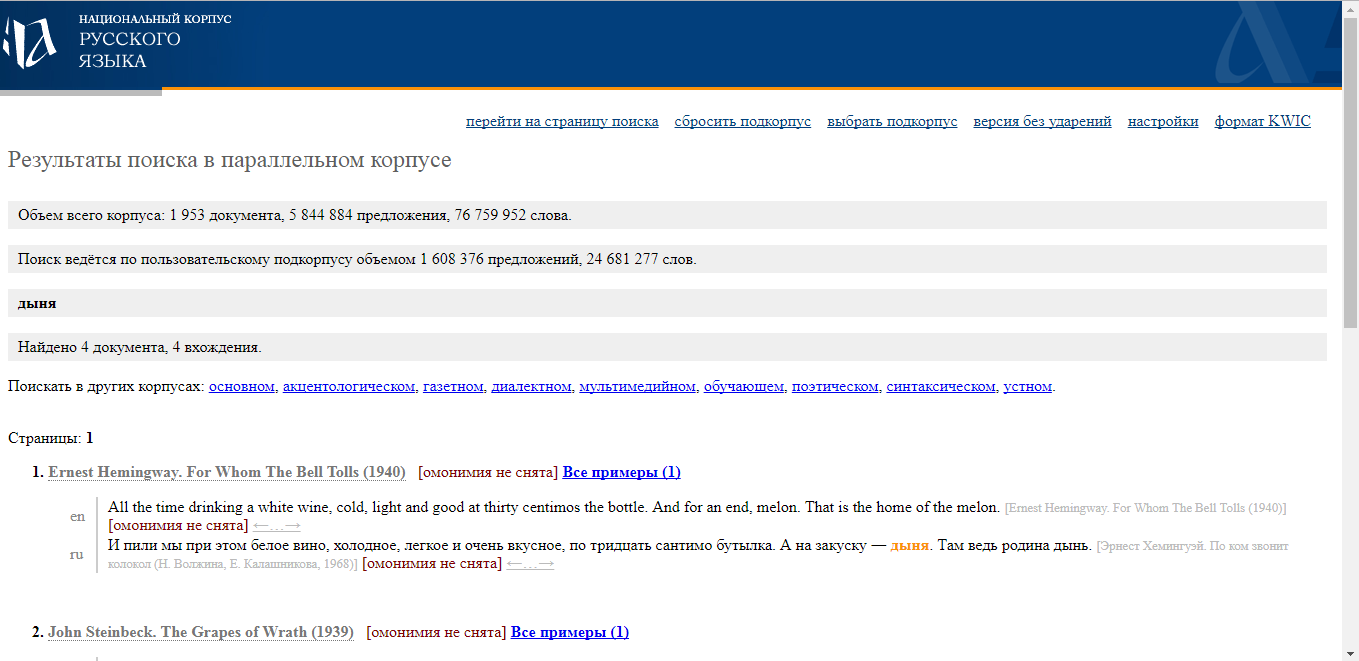
Чтобы до конца в этом убедиться, я использовала две формулы (меры разброса моделей перевода): «отношение абсолютной частоты самой частотной модели перевода к количеству различных моделей» и «отношение абсолютной частоты самой частотной модели перевода к частоте второй». Оценивать конечные данные после подсчёта некорректно, так как статистика в НКРЯ для данных слов неполная, а вот проанализировать саму формулу и соотношение будет логичнее.

* Первая формула рассчитывалась так: 12 к 37. Данное соотношение показывает, что на каждую модель перевода в среднем приходится сравнительно немного контекстов, а самая частотная из всех моделей занимает небольшой процент (примерно 32%) от общего числа возможных соответствий, ее частота практически не отличается от частот остальных.
* Вторая мера разброса рассчитывалась по-другому: 12 к 11. Можно увидеть, что разница между частотой самой частотной модели перевода и второй модели едва заметна, отличается буквально на один пример. Это присуще лингвоспецифичному слову, так как у него первые несколько моделей перевода должны употребляться практически одинаковое количество раз.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Слово** | **Количество вхождений слова** | **Количество моделей перевода** | **Самая частотная модель перевода** | **Вторая модель по частоте перевода** |
| Слово "злобный" | 72 | 37 | 12 | 11 |
| Слово "дыня" | 4 | 2 | 4 | 1 |

|  |  |  |
| --- | --- | --- |
| **Формулы вычисления** | **Слово "злобный"** | **Слово "дыня"** |
| (F (Mmax))/(NumM) | 12/37 | 4/2 |
| (F (Mmax)/F (Msec)) | 12/11 | 4/1 |

**2)** Далее рассмотрим второе слово: «дыня». На сайте показано, что слово имеет всего 4 вхождения и маленькое количество примеров. После подсчётов оказалось, что моделей перевода всего 2, самая частотная из них употребляется 4 раза («melon»), а вторая по частоте лишь 1 раз («cantaloupe»). В этом случае также можно заметить, что наблюдается маленькое число моделей перевода, а это черта неспецифичного слова.



* Что же касается мер разброса, то по первой формуле мы видим соотношение: 4/2. То есть самая частотная из всех моделей занимает огромный процент от общего числа возможных соответствий, чего никак не может быть у лингвоспецифичного слова.
* Вторая мера разброса также доказывает, что это слово неспецифичное. В этой формуле мы видим соотношение: 4/1. Из этого следует, что частота самой частотной модели перевода намного отличается от частоты второй модели (аж на три примера, а не как у слова «злобный»). Такого не может наблюдаться у лингвоспецифичного слова.

**Выводы:**

* Выдвинутые предположения и гипотезы действительно оправдались. В этом помог не только сайт НКРЯ (своими моделями перевода, примера, контекстами, частотами), но и меры разброса моделей перевода, которые давались в самом ДЗ7.
* Слово «злобный» действительно является лингвоспецифичным, это выяснилось благодаря огромному количеству моделей перевода на английском языке, а также по соотношениям, полученным по формулам (мерам разброса).
* Слово «дыня» оказалось неспецифичным, как, впрочем, и предполагалось. Маленькое количество моделей перевода, абсолютно противоположные показатели по формулам это с точностью доказали.
* НКРЯ оказался очень полезным сайтом, который способен не только выдать огромное количество различных примеров со словами на многих языках и показать большой список моделей перевода, но также разобраться в лингвоспецифичности слов, что безусловно полезно для лингвистов.