

## *Использование параллельного корпуса для количественного изучения лингвоспецифичной лексики*

Мне нужно было выбрать два слова, гипотетические лингвоспецифичное и неспецифичное. Я взяла два слова.

Начнем с того слова, которое я охарактеризовала как «лингвоспецифичное», а именно потому, что, проанализировав все страницы, которых было 3, и все вхождения, которых оказалось 52, обнаружила 18 различных моделей перевода. На русском языке слово обозначает прилагательное «обширный». В последующем я отфильтровала результаты от большего к меньшему и моделью, имеющей наибольшую частоту вхождений, оказалось слово «wide», его количество раз равно 9. Следующим по количеству вхождений, равному 6, оказалось слово «vast».

- Дальше я вычислила отношение абсолютной частоты самой частотной модели перевода к количеству различных моделей. Я это сделала следующим образом: выделила ячейку, где показано наибольшее количество вхождений (9 раз), (E2) и поделила на общее количество моделей перевода, которых 18.
- В последующем шаге я вычислила среднюю частоту вхождений на одну модель, для этого я вычислила сумму всех вхождений (E2-E19) и разделила на количество моделей, количество которых равно 18.
- Дальше я вычислила отношение абсолютной частоты самой частотной модели перевода к частоте второй, для этого количество вхождений самой частотной модели перевода (wide(E2)) были поделены на количество вхождений второй по очереди модели перевода (vast(E3)).
- После этого я вычислила отношение абсолютной частоты самой частотной модели перевода к общему количеству вхождений. Для этого я количество вхождений самой частотной модели перевода (wide(E2)) разделила на сумму всех вхождений моделей перевода (E2-E19).

А теперь обратим свое внимание на слово, которое неспецифичное. На русском языке я взяла слово «любовь», при переводе на английский язык, у меня оказалось очень большое количество вхождений, поэтому я проанализировала первые две первые страницы и нашла всего две модели перевода, а именно «love», «affection». Наиболее встречающейся моделью перевода, что неудивительно, оказалось «love». С этими двумя словами я проделала те же самые вычисления, что и с лингвоспецифичными. А именно:

- Отношение абсолютной частоты самой частотной модели перевода к количеству различных моделей. Я это сделала следующим образом: выделила ячейку, где показано наибольшее количество вхождений ((74 раза)B2) и поделила на общее количество моделей перевода, которых 2.
- Средняя частота вхождений на одну модель - для этого я вычислила сумму всех вхождений (B2-B3) и разделила на количество моделей, количество которых равно 2.
- Абсолютная частота самой частотной модели перевода к частоте второй, для этого я поделила количество вхождений самой частотной модели перевода (love (B2)) на количество вхождений второй по очереди модели перевода (affection (B3)).
- Отношение абсолютной частоты самой частотной модели перевода к общему количеству вхождений. Для этого я количество вхождений самой частотной модели перевода (love (B2)) разделила на сумму всех вхождений моделей перевода (B2-B3).

Из этого мы можем сделать следующие выводы. Лингвоспецифичное слово имеет большое количество моделей перевода (у меня их 18), в отличие от неспецифичного, которое имеет всего 2 модели. Помимо этого отношение абсолютной частоты самой частотной модели перевода к количеству различных моделей у неспецифичного слова у меня получилось 37, в то время как у специфичного 0,5. Также средняя частота вхождений на одну модель у неспецифичного у меня получилась 39,5, а у специфичного – 2,3. Абсолютная частота самой частотной модели перевода к частоте второй у неспецифичного слова (моделей которого всего 2) равна 14,8, а у специфичного –

1,5. И последнее вычисление, которое имеет значение, это отношение абсолютной частоты самой частотной модели перевода к общему количеству вхождений, где у неспецифичного слова получилось – 0,936, а у специфичного – 0,214. Все эти числа показывают наглядную разницу между лингвоспецифичным словом и неспецифичным. Так как у лингвоспецифичного слова самая частотная из всех моделей занимает небольшой процент от общего числа возможных соответствий, ее частота несильно отличается от частот остальных, в отличие от неспецифичного, где все в точности наоборот и это четко видно на цифрах, которые указаны выше в соответствии с вычислениями.