

# Использование параллельного корпуса НКРЯ для изучения лингвоспецифичной лексики

## Часть 1

Гипотетическое лингвоспецифичное слово «простокваша». Данное слово мне кажется более характерным для русского языка и культуры в целом, поэтому я предполагаю, что оно может считаться лингвоспецифичным.

В параллельном англо-русском корпусе НКРЯ было найдено 10 документов, 16 вхождений.

Модели перевода:

1. Prostokvasha (1)
2. Curd(s) and whey (7)
3. Curdled milk (1)
4. Yogurt (1)
5. Junket (4)
6. Whey (1)

отношение абсолютной частоты самой частотной модели перевода ( $F(M_{max})$ ) к количеству различных моделей ( $NumM$ ) – 1,16

средняя частота вхождений на одну модель ( $F(O)/NumM$ , где  $F(O)$  — общее количество вхождений) – 2,6

отношение абсолютной частоты самой частотной модели перевода к частоте второй ( $F(M_{max})/F(M_{sec})$ ) -1,75

отношение абсолютной частоты самой частотной модели перевода к общему количеству вхождений ( $F(M_{max})/F(O)$ ) -0,437

$F(O)$	16	$(F(M_{max}))/NumM$	1,16
$F(M_{max})$	7	$(F(O))/NumM$	2,6
$NumM$	6	$(F(M_{max})/F(M_{sec}))$	1,75
$F(M_{sec})$	4	$(F(M_{max})/F(O))$	0,437

Вывод:

Предполагается, что у «лингвоспецифичного» слова должен наблюдаться большой список возможных моделей перевода, и на каждую будет в среднем приходится сравнительно немного контекстов, а самая частотная из всех моделей будет занимать небольшой процент от общего числа возможных соответствий, ее частота не будет сильно отличаться от частот остальных. Русское слово «простокваша», основываясь на результатах анализа, может быть названо лингвоспецифичным, но с трудностью, так как с одной стороны слово имеет несколько отдаленно похожих моделей перевода, а с другой стороны частота самой частотной модели относительно сильно отличается от других моделей.

## Часть 2

Гипотетически не лингвоспецифичное слово «ученик».

Как вариант гипотетически нелингвоспецифичного слова, я выбрала «ученика». На мой взгляд данное слово встречается как в английском, так и в русском языке довольно часто. И в двух языках не обладают какими-либо специфичными характеристиками.

В англо-русском параллельном корпусе было найдено 40 документов, 65 вхождений слова «ученик».

Модели перевода:

1. «pupil» (27)
2. «disciple» (13)
3. «student» (9)
4. «head boy» (2)
5. «boy» «schoolboy» (7)
6. «learner» (2)
7. «scholar» (1)
8. «prentice» (1)

Значения:

$F(O)$ - 65

$F(M_{max})$ - 27

$NumM$ - 8

$F(M_{sec})$ - 13

отношение абсолютной частоты самой частотной модели перевода ( $F(M_{max})$ ) к количеству различных моделей ( $NumM$ )-3,375

средняя частота вхождений на одну модель ( $F(O)/NumM$ )- 8,125

отношение абсолютной частоты самой частотной модели перевода к частоте второй ( $F(M_{max})/F(M_{sec})$ )- 2,06

отношение абсолютной частоты самой частотной модели перевода к общему количеству вхождений ( $F(M_{max})/F(O)$ )-0,414

$F(O)$	65	$(F(M_{max}))/NumM$	3,375
$F(M_{max})$	27	$(F(O))/NumM$	8,125
$NumM$	8	$(F(M_{max})/F(M_{sec}))$	2,06
$F(M_{sec})$	13	$(F(M_{max})/F(O))$	0,414

Вывод:

Слово «ученик» с трудом может быть названо лингвоспецифичным, так как на каждую модель перевода приходится сравнительно большое количество контекстов. Частота самой частотной модели значительно превосходит другие модели. И так как лингвоспецифическое слово – это слова, у которых эквивалент в другом языке вовсе отсутствует, и в случае необходимости

приходится прибегать к заимствованию слова или использовать описательное определение, то слово «ученик» будет являться нелингвоспецифичным.