

Изучение лингвоспецифичности слов «захолустье» и «деревня»

Для данного задания я выбрал слова «захолустье» и «деревня». По моим предположениям, слово «захолустье» должно было являться лингвоспецифичным для русского языка, а «деревня» - нет. Согласно информации, данной в задании «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.» Используя параллельный корпус НКРЯ для сравнения моделей перевода этих слов с русского на английский, я получил следующие результаты, которые подтвердили мою гипотезу:

Слово	F(Mmax) - абсолютная частота самой частотной модели перевода	F(Msec) - вторая по частоте модель перевода	F(O) - частотность слова в оригинале	NumM - количество различных эквивалентов	F(Mmax)/NumM	F(O)/NumM	F(Mmax)/F(Msec)	F(Mmax)/F(O)
захолустье	1	1	8	7	0,14	1,14	1,00	0,13
деревня	212	128	563	3	70,67	187,67	1,66	0,38

Захолустье	
Модель перевода	Частота
Godforsaken corner	1
A cluster of villages	1
Foreign country	1
The wilds	1
Out-of-the-way place	1
Little town	1
Peaceful corner	1
Не переведено	1

Деревня	
Модель перевода	Частота
Village	212
Country, country side, country place и т.п.	128
Town	37
Не переведено/слито с другими контекстами (например, "поместье в деревне" = просто "estate")	186

Можно заметить, что у слова «захолустье» наблюдается большое количество моделей перевода: на 8 вхождений 7 различных эквивалентов (например, «Godforsaken corner», «a cluster of villages», «foreign country» и т.д.; при этом еще один раз слово не было переведено на английский вообще). Поэтому отношение контекстов употребления в оригинале (частотности) к количеству этих моделей (F(O)/NumM) стремится к своему минимуму – единице, т.е. на каждую модель перевода приходится практически минимальное количество контекстов. Отношение частоты самой употребляемой модели перевода ко второй по употребляемости модели (F(Mmax)/F(Msec)) также равно единице, т.е. ее частота употребления почти не отличается от частот остальных. Наконец отношение частоты этой модели к количеству различных эквивалентов (F(Mmax)/NumM) очень низко, а занимает она совсем небольшой процент от общего числа возможных соответствий (F(Mmax)/F(O) = 0,13 = 13%), что также является признаком лингвоспецифичного слова.

У слова «деревня» количество моделей перевода меньше, а частота его употребления в оригинале значительно больше, чем у слова «захолустье», поэтому и отношение контекстов употребления к количеству моделей перевода ($F(O)/NumM$), как можно заметить, намного выше, т.е. на каждую модель приходится в среднем сравнительно большое количество контекстов. Что касается самых частотных моделей перевода, то слово «деревня» чаще всего переводится на английский как «village». Есть и вторая довольно употребительная модель: «country», «country side» и т.п., но встречается она реже; поэтому отношение частотности этих двух моделей ($F(M_{max})/F(M_{sec})$) выше, чем у слова «захолустье», хотя тотального превосходства первой модели над второй все же не наблюдается. Однако они являются основными, и количество моделей перевода намного меньше, чем в первом случае, на основании чего можно судить о неспецифичности этого слова. Также отношение самой частотной модели перевода к количеству различных эквивалентов сравнительно высоко ($F(M_{max})/NumM$), а занимает она большую (тоже сравнительно) часть от общего числа возможных соответствий ($F(M_{max})/F(O) = 0,38 = 38\%$), что не соответствует признакам лингвоспецифичного слова.

Таким образом, основываясь на признаках лингвоспецифичных слов и данных, полученных из НКРЯ, я пришел к выводу, что слово «захолустье» действительно является лингвоспецифичным, а слово «деревня» таковым не является.