

Домашнее задание № 7

Использование параллельного корпуса для количественного изучения лингвоспецифичной лексики

Для исследования я беру два слова «белобрысый» и «рыжий». Не проводив подсчёты, я предположила, что слово «белобрысый» является лингвоспецифичным словом, то есть оно не имеет чёткого аналога в английском языке. Используя параллельный корпус НКРЯ, я провела подсчеты таких критериев как отношения абсолютной частоты самой частотной модели перевода (NumM) средняя частота вхождений на одну модель ($F(O)/NumM$), отношение абсолютной частоты самой частотной модели перевода к частоте второй ($(Mmax)/F(Msec)$), отношение абсолютной частоты самой частотной модели перевода к общему количеству вхождений ($(F(Mmax)/F(O))$).

Лингвоспецифичное слово «белобрысый»

Переводы на английский: gray-haired(1), bland(1), blond(1), fair-haired(2), tow-haired(1), white(1)

[6 различных переводов слова]

Найдено 8 документов, 10 вхождений. В одном из текстов слово было упущено, в другом это слово употребляется два раза в одном тексте, поэтому при подсчетах учитываются 7 документов и 7 вхождений.

$$NumM = 0,33$$

$$F(O)/NumM = 21,2$$

$$(Mmax)/F(Msec) = 2$$

$$(F(Mmax)/F(O)) = 0,29$$

Неспецифичное слово «рыжий»

Переводы на английский: red-headed(11) with red hair(7) red-haired(22) red(13)reddy-brown(1) tawny(1) ginger-haired(1) ginger(1)brown(3) yellow-brown(1) pink of hair(1) carroty(1) fair-headed(1)tan-coloured (1) chestnut(1) sandy-haired(1).

Объединяем слова с одним корнем «red» и получается 54 употребления (но самым распространенным словом среди этих слов является «red-haired»), объединяем также однокоренные слова «ginger» и «brown».

Red(54), brown(4) ginger(2), tawny(1), pink of hair(1), carroty(1), fair-headed(1), tan-coloured (1), chestnut(1), sandy-haired(1). [10 различных переводов слова]

Найдено 57 документов, 187 вхождений.

$$NumM = 5,4$$

$$F(O)/NumM = 34,6$$

$$(Mmax)/F(Msec) = 13,5$$

$$(F(Mmax)/F(O)) = 0,29$$

Опираясь на подсчёты, можно утверждать, что слово «белобрысый» является лингвоспецифичным, так как не имеет основного аналога в английском языке. У него наблюдается большой список моделей перевода, а самая частотная из всех моделей занимает небольшой процент от общего числа возможных соответствий, и ее частота несильно отличается от частот остальных (2 к 1). В то время как у слова «рыжий» легко находится основная модель перевода с частотой употребления 54 раза, в то время как следующее по частоте слово употребляется 4. У неспецифичного слова отношение абсолютной частоты самой частотной модели перевода значительно выше, чем у лингвоспецифичного слова, как и остальные показатели в целом, такие как средняя частота вхождений на одну модель, отношение абсолютной частоты самой частотной

	"белобрысый"	"рыжий"
NumM	0,33	5,4
F (O)/NumM	21,2	34,6
(Mmax)/F (Msec)	2	13,5
(F (Mmax))/F (O)	0,29	0,29

модели перевода к частоте второй, отношение абсолютной частоты самой частотной модели перевода к общему количеству вхождений.