

Отчет по домашнему заданию №7 «Использование параллельного корпуса для
количественного изучения лингвоспецифичной лексики»

«Удаль» (английский параллельный корпус, 21 вхождение)

rakishness	1
abandonment	1
assurance	1
defiant	1
daring	2
violence	1
effrontery	1
reckless (jockeying)	2
feats	1
prowess	2
fearless confidence	1
audacity	1
bravado	1
bravery	1

«Платок» (английский параллельный корпус, 1014 вхождений)

(hand)(neck)(breast)kerchief	597
flash of white cotton	1
tissue	3
(face) flannel	2
scarf	14
babushka	1
linen	1
cloth	2
shawl	37
bandan(n)a	7
rag	3
wrap(ping)	2
bandage	1
sheet	1
muffler	4
(pocket)napkin	4
gravatt	1
ribbon	1
apron	1
sling	1
skirt	1
dresses	1
gown	1

Расчеты по формулам мер разброса для слова «удаль»:

- отношение абсолютной частоты самой частотной модели перевода ($F(M_{\max})$) к количеству различных эквивалентов ($\text{Num}M$):

Самые частотные модели перевода – *daring, reckless, prowess* – имеют одинаковую частоту (2) – это F (Mmax). NumM = 13 (не включая одну из самых частотных моделей).

Для расчетов возьмем одну из трех моделей, например, *daring*:

$$2/13 = 0.15$$

- средняя частота вхождений на один эквивалент (F (O)/NumM, где F (O) — общее количество вхождений):

$$F (O) = 21; \text{NumM} = 14$$

$$21/14 = 1.5$$

- отношение абсолютной частоты самой частотной модели перевода к частоте второй (F (Mmax)/F (Msec)):

Самые частотные модели перевода – *daring, reckless, prowess* – имеют одинаковую частоту (2) – это F (Mmax). F (Msec) – все остальные эквиваленты, т.к. они все имеют одинаковую частоту (1).

$$2/1 = 2$$

- отношение абсолютной частоты самой частотной модели перевода к общему количеству вхождений (F (Mmax)/F (O)):

$$F (Mmax) = 2; F (O) = 21$$

$$2/21 = 0.09$$

Расчеты по формулам мер разброса для слова «платок»:

- отношение абсолютной частоты самой частотной модели перевода (F (Mmax)) к количеству различных эквивалентов (NumM):

Самая частотная модель перевода - *(hand)(neck)(breast)kerchief* (597) – это F (Mmax). NumM = 22 (не включая самую частотную модель).

$$597/22 = 27.1$$

- средняя частота вхождений на один эквивалент (F (O)/NumM, где F (O) — общее количество вхождений):

$$F (O) = 1014; \text{NumM} = 23$$

$$1014/23 = 44$$

- отношение абсолютной частоты самой частотной модели перевода к частоте второй (F (Mmax)/F (Msec)):

$$F (Mmax) = 597 ((hand)(neck)(breast)kerchief); F (Msec) = 37 (shawl)$$

$$597/37 = 16.1$$

- отношение абсолютной частоты самой частотной модели перевода к общему количеству вхождений (F (Mmax)/F (O)):

$$F (Mmax) = 597; F (O) = 1014$$

$$597/1014 = 0.58 \sim 0.6$$

ВЫВОД: Таким образом, слово «удаль» действительно можно назвать лингвоспецифичным, поскольку у него большой список возможных моделей перевода, все они недостаточно точны, а частота самых популярных моделей практически не отличается от менее популярных (2 против 1).

Если говорить о слове «платок», то его можно назвать неспецифичным, даже несмотря на множество вариаций перевода, поскольку, если сравнивать его по расчетам со словом «удаль», частота самой популярной модели сильно отличается от остальных, даже от второй по частоте (597 против 37). У него также, в отличие от «удали», существует множество контекстов (носовой платок, шейный платок, теплая шаль и т.п.).