

Задание 7

Для проведения исследования были выбраны 2 слова: самоотверженность (предположительно лингвоспецифичное) и доблесть (предположительно нелингвоспецифичное). Поиск будет проводится по параллельному русско-английскому корпусу объемом 1 608 376 предложений, 24 681 277 слов.

Начнем с предположительно лингвоспецифичного слова 'самоотверженность':

Общее количество вхождений: 16 (12 документов), список моделей перевода:

- 1) Mythology
- 2) Valor
- 3) Selfdenial x3
- 4) Self-abnegation
- 5) Unselfishness x2
- 6) Self-sacrifice
- 7) Self devotion x2
- 8) Magnanimous things
- 9) Complaisance
- 10) Disinterestedness

Занесём данные в таблицу:

Модель перевода (самоотверженность)	Количество раз
Mythology	1
Valor	1
Selfdenial	3
Self-abnegation	1
Unselfishness	2
Self-sacrifice	1
Self devotion	2
Magnanimous things	1
Complaisance	1
Disinterestedness	1

Посчитаем по предложенным формулам:

- 1) Отношение абсолютной частоты самой частотной модели перевода к количеству эквивалентов $F(M_{\max})/NumM = 3/9 = 0,33$
- 2) Среднее число вхождений на 1 эквивалент $F(O)/NumM = 16 \text{ (число вхождений)} / 10 \text{ (варианты перевода)} = 1,6$
- 3) Отношение абсолютной частоты самой частотной модели ко второй по частотности $F(M_{\max})/F(M_{\text{sec}}) = 3/2 = 1,5$
- 4) Отношение абсолютной частоты самой частотной модели к общему числу вхождений $F(M_{\max})/F(O) = 3/16 = 0,188$

Перейдём к предположительно нелингвоспецифичному слову 'доблесть':

Общее количество вхождений: 74 (29 документов), список моделей перевода:

- 1) Daring
- 2) Valor / Valour x11
- 3) Courage x4
- 4) Dignity x3
- 5) [is] meritorious
- 6) Virtue x6
- 7) Glory x4
- 8) Gallantry
- 9) Flaunt one's strength
- 10) Prowess x3
- 11) Heroism
- 12) Bravery x2
- 13) Gallant manner
- 14) Honour

Занесём данные в таблицу:

Модель перевода (доблесть)	Количество раз
Daring	1
Valor x11	11
Courage x4	4
Dignity x3	3
[is] meritorious	1
Virtue x6	6
Glory x4	4
Gallantry	1
Flaunt one's strength	1
Prowess x3	3
Heroism	1
Bravery x2	2
Gallant manner	1
Honour	1

Посчитаем по предложенным формулам:

- 1) Отношение абсолютной частоты самой частотной модели перевода к количеству эквивалентов $F(M_{\max})/NumM = 11/14 = 0,786$
- 2) Среднее число вхождений на 1 эквивалент $F(O)/NumM = 16 \text{ (число вхождений)} / 10 \text{ (варианты перевода)} = 74/14 = 5,286$
- 3) Отношение абсолютной частоты самой частотной модели ко второй по частотности $F(M_{\max})/F(M_{\text{sec}}) = 11/6 = 1,833$
- 4) Отношение абсолютной частоты самой частотной модели к общему числу вхождений $F(M_{\max})/F(O) = 11/74 = 0,149$

Проанализируем итоги:

Для первого слова индекс Герфиндаля = 0,188, для второго 0,149. Учитывая, что максимум = 1, эти значения довольно малы. Это особенно заметно при сравнении минимально возможного индекса Герфиндаля для этих слов – для первого слова он равен $1/9 = 0,11$ для второго слова он равен $1/14 = 0,071$. Сравнив полученные значения индексов с потенциальными минимумами, получаем, что индекс Герфиндаля для первого слова практически минимально возможный, а значит, оно действительно лингвоспецифично. Значение индекса второго, в свою очередь, сравнительно несколько выше, а значит, степень его лингвоспецифичности меньше.

Действительно, моделей перевода для первого слова 10 при 16 вхождениях, а самый часто встречающийся вариант появился 3 раза, когда следующий за ним – 2. А значит, большого разрыва между ними нет. Следовательно, слово 'самоотверженность' лингвоспецифично.

Слово 'доблесть', напротив, при 74 вхождениях имеет всего 15 моделей перевода, а самая популярная модель встречается практически в два раза чаще, чем следующая за ней. Вкупе со сравнительно высоким индексом Герфиндаля это слово можно отнести к нелингвоспецифичным.

Вывод: слово 'самоотверженность' лингвоспецифично, слово 'доблесть' нелингвоспецифично.