

Использование параллельного корпуса для количественного изучения лингвоспецифичной лексики

Для изучения лингвоспецифичной лексики мною выбран украинский параллельный корпус в НКРЯ. Будет сравниваться перевод слов с украинского на русский язык.

Гипотеза: допустим, слово 'miteць' является неспецифичным, а слово 'плахта' – лингвоспецифичным. Одинаковый перевод в одном и том же документе не учитывается.

1. 'Митець'

Найдено 12 документов, 18 вхождений.

С украинского на русский 'Митець' (оригинал)	
Вариант перевода	Частота употребления
Мастер	6
Художник	2
Мудрец	1
Мастак	1
Проворный	1
С русского на украинский 'Митець' (перевод)	
Вариант перевода	Частота употребления
Мастер	1
Искусник	1

Было найдено 11 документов с переводом с украинского на русский. Полученные результаты можно наблюдать в таблице.

Был найден 1 документ, где был осуществлен перевод с русского на украинский – Н. В. Гоголь. Мертвые души (1835-1852). В данном случае было 2 разных русских слова, которые были переведены на 'miteць' – мастер, искусник. Эти данные не будут учитываться, так как не являются оригиналом (украинский язык).

С учетом этого имеем 11 документов с 16 вхождениями.

- отношение абсолютной частоты самой частотной модели перевода к количеству различных моделей (NumM) – 1,2 $(F(M_{max}) / \text{NumM})$
- средняя частота вхождений на одну модель $(F(O) / \text{NumM})$ – 3,2
- отношение абсолютной частоты самой частотной модели перевода к частоте второй $(F(M_{max}) / F(M_{sec}))$ – 3
- отношение абсолютной частоты самой частотной модели перевода к общему количеству вхождений $(F(M_{max}) / F(O))$ – 0,375

В 54,5% случаев это слово переводится как «мастер». Контекст везде одинаковый – слово используется для обозначения человека, который сделал, придумал, создал что-то необычное, уникальное, например в искусстве.

В таком случае слово «miteць» не ведет себя похоже на лингвоспецифичное слово из-за отсутствия небольшой разницы вариантами перевода на русский язык, но по результатам стимулов русского перевода является общим термином для многих русских наименований.

Вывод: украинское слово «miteць» является неспецифичным – гипотеза доказана. Однако представляет собой общий термин для таких русских слов, как «мастер», «художник», «мудрец», «мастак», «проворный».

2. 'Плахта'

Найдено 14 документов, 33 вхождения.

С украинского на русский 'Митець' (оригинал)	
Вариант перевода	Частота употребления
Плахта	10
Панева	1
Полотнище	1
С русского на украинский 'Плахта' (перевод)	
Вариант перевода	Частота употребления
Плахта	1

Было найдено 12 документов с переводом с украинского на русский. Полученные результаты можно наблюдать в таблице.

Было найдено 2 документа, где был осуществлен перевод с русского на украинский – Н. В. Гоголь. Вий (1835) и Пантелеймон Кулиш. Черная рада, Хроника 1663 года (1846-1857). Однако во втором документе отсутствует перевод на украинский (перевода текста полностью отсутствует). Эти данные не будут учитываться, так как не являются оригиналом (украинский язык), так как ожидается, что лингвоспецифичная лексика скорее появится при создании оригинального текста, чем при переводе.

С учетом этого имеем 12 документов с 31 вхождением.

- отношение абсолютной частоты самой частотной модели перевода $(F(M_{max}))$ к количеству различных моделей $(NumM)$ – 3,3
- средняя частота вхождений на одну модель $(F(O)/NumM)$ – 10,3
- отношение абсолютной частоты самой частотной модели перевода к частоте второй $(F(M_{max})/F(M_{sec}))$ – 10
- отношение абсолютной частоты самой частотной модели перевода к общему количеству вхождений $(F(M_{max})/F(O))$ – 0,323

В 83% случаев это слово переводится как «плахта», что соответствует оригинальному слову. Причина этого значение слова «плахта» - кустарная украинская ткань, полосатая или клетчатая, а также четырёхугольный отрез её, носимый вместо юбки. Стоит учесть, что во всех случаях «плахта» употреблялась в значении четырёхугольного отреза, юбки, что было видно из контекста. Однако 1 раз слово имело значение именно ткани и в этом случае переводилось как «полотно»:

До хати всотується біла **плахта**, всотується без кінця й міри. [Василь Стефаник. Скін (1901)]

В хату всовывається біле **полотнище** и летит, летит, — нет ему конца и края.

В таком случае слово «плахта» не ведет себя похоже на лингвоспецифичное слово из-за отсутствия вариантов перевода на русский язык. Но так как это является исключительно частью украинского костюма или тканью, которая производилась на украинских землях, можно предположить, что данное слово вовсе не используется в русском языке, ведь 2 документа с переводом с русского на украинский, о которых говорилось выше, описывают Украину – произведение Гоголя и Кулиша. Поэтому можно утверждать, что слово «плахта» лингвоспецифичное.

Вывод: украинское слово «митець» является лингвоспецифичным – гипотеза доказана. Это связано с уникальностью и полным отсутствием использования в русском языке (только для перевода в контексте описания Украины).