**Отчёт по работе с параллельными корпусами НКРЯ.**

**Задачи.**

* Выбрать два слова, гипотетически лингвоспецифичное и неспецифичное.
* Используя один из параллельных корпусов НКРЯ и меры разброса моделей перевода, подтвердить свои гипотезы.

**Ход работы.**

* Мною были выбраны два слова: «авось» и «собака». Моя гипотеза: слово «авось» - лингвоспецифичное (имеет большое количество моделей перевода), а слово «собака» таковым не является.
* Далее, используя параллельный корпус НКРЯ (английский) я произвела поиск и затем исследовала полученные примеры на предмет вариативности моделей перевода (сначала для слова «авось», затем для слова «собака»). Количество страниц с примерами перевода первого слова – 4 штуки. Поэтому для второго я сочла нужным анализировать исключительно первые 4 страницы с примерами.
* Благодаря анализу примеров в параллельном корпусе мне стали известны возможные модели перевода, а также количество вхождений каждой отдельной модели.
* Данные я перенесла в таблицу Excel для дальнейших подсчётов, необходимых для доказательства лингвоспецифичности одного и нелингвоспецифичности другого слова. (Таблицу прилагаю к Репозиторию “hw7”).
* Оценить степень специфичности того или иного слова можно с помощью нескольких мер разброса моделей перевода, с помощью которых возможно узнать:
  + - отношение абсолютной частоты самой частотной модели перевода (F (Mmax)) к количеству различных моделей (NumM);
    - среднюю частоту вхождений на одну модель (F (O)/NumM, где F (O) - общее количество вхождений);
    - отношение абсолютной частоты самой частотной модели перевода к частоте второй (F (Mmax)/F (Msec));
    - отношение абсолютной частоты самой частотной модели перевода к общему количеству вхождений (F (Mmax)/F (O)).
* В таблице Excel я произвела сортировку (от самой частоупотребляемой модели до менее употребляемой) и подсчёты, согласно вышеупомянутым формулам.
* На основе результатов проведённого исследования можно сделать вывод: действительно, «авось» является лингвоспецифичным словом, так как имеет довольно много моделей перевода, на каждую из которых приходится сравнительно немного контекстов, а частота самой частотной из них не сильно отличается от остальных. В свою очередь, слово «собака» обладает абсолютно противоположной характеристикой: оно имеет одну основную модель перевода, а другие модели, если и попадаются, то крайне редко. Значит, моя гипотеза верна, и данное слово - нелингвоспецифичное.