

Научный семинар Школы лингвистики НИУ ВШЭ

с участием лингвистической лаборатории по корпусным
технологиям и лаборатории языков Кавказа

Любовь Нестеренко

Школа лингвистики НИУ ВШЭ

Использование мультязычных параллельных корпусов для квантитативных исследований грамматики (на примере залоговых конструкций)

Мультязычные параллельные корпуса открывают возможности для применения в типологии квантитативных методов, таких, как кластерный анализ, сетевой анализ, машинное обучение и т.д. В докладе будет продемонстрировано, как мультязычные параллельные корпуса и машинное обучение могут быть использованы для исследования залоговых конструкций. В рамках эксперимента рассматривается противопоставление «актив vs. пассив с выраженным агенсом». Для определения того, чем может быть обусловлен выбор между активом и пассивом, были построены модели логистической регрессии. Список признаков для обучения включает в себя одушевленность участников ситуации, их семантические роли, наличие упоминания в предшествующем контексте и др. Модели рассматриваемых языков имеют схожие значимые признаки, однако наборы признаков, оказавшихся значимыми, совпадают не полностью. В качестве материала используется корпус, включающий в себя книги о Гарри Поттере на девяти европейских языках.

Старая Басманная 21/4,
аудитория 518
5 апреля (пятница) в 10:30

Сайт семинара:
ling.hse.ru/SciSeminar

