**Доклад “Multi-level student essay feedback in a learner corpus”**

О.И. Виноградова, О.Н. Ляшевская, И.М. Пантелеева

12.06.2017

В данном докладе показано применение компьютерных инструментов оценки лексического, грамматического и синтаксического уровней текста с помощью автоматической и полуавтоматической разметки студенческих сочинений в обучающем корпусе REALEC. Сравнение проводилось на выборке размером примерно 2000 эссе. Первым этапом работы являлся синтаксический парсинг текста с помощью инструмента UDPipe, с помощью которого было исследовано два критерия: средняя глубина деревьев и количсетво придаточных предложений. На втором этапе была проанализирована выдача лексического инспектора. На третьем – исследовалась аннотация к текстам. Благодаря этому исследованию предполагалось выявить те критерии, которые разнятся для ‘лучших’ и ‘худших’ эссе.

Результаты работы следующие:

* Средняя глубина дерева не является важным критерием для оценки синтаксической сложности текста
* Количество придаточных предложений в ‘лучших’ сочинениях больше, чем в ‘худших’
* Длина слов и количество повторении являются незначительными критериями для проверки уровня владения языком
* Статистика, связанная с автоматическим и ручным оцениванием работ