**Автоматическая генерация вопросов типа Multiple Choice на основе данных учебного корпуса**

Выполнил студент 2 курса бакалавриата направления

«Фундаментальная и компьютерная лингвистика»

Логин Никита Вячеславович

Для научно-учебной группы «REALEC для реально необходимых слов»

Системы BRAT, используемая в учебном корпусе REALEC, хранит аннотации в открытом формате, что обеспечивает возможность создания лексико-грамматических упражнений экзаменационного типа на основе реальных ошибок студентов. Задания экзамена IELTS делятся на 4 категории: “Open-close”, “Word form”, “Short answer” и “Multiple Choice”. Ранее участниками научно-учебной группы уже было реализовано программное решение (<https://github.com/kategerasimenko/realec-exercises/blob/master/realec_grammar_exercises.py>), позволяющее генерировать вопросы первых трёх типов, моей задачей было добавить возможность порождения заданий четвёртого типа.

Для решения поставленной задачи я воспользовался собранным участниками группами словарём форм английского языка, сохранённым в формате json. В виде программного кода удалось воплотить 6 из 10 предложенных решений для различных контекстов ошибки. Исходный код полученного решения доступен по адресу: <https://github.com/nicklogin/realec_multiple_choice/realec_grammar_exercises/py>