

РЕЦЕНЗИЯ
научного руководителя
на выпускную квалификационную работу обучающегося СПбГУ
Эгервари Кристиана Эрика
по теме «Выравнивание русских предложных конструкций и их эквивалентов в агглютинативных языках при создании параллельных корпусов»


Работа Кристиана Эгервари посвящена разработке методов и инструментов выравнивания русских предложных конструкций и их эквивалентов в русском и венгерском языках. Выравнивание текстов (англ. alignment) - необходимый этап при создании и аннотировании параллельных корпусов. Он также находит место в задачах составления двуязычных словарей, извлечения двуязычной лексики, статистического машинного перевода, обучения лингвистических моделей, проведении компаративистских исследований. Идеальным является пословное выравнивание, но оно часто оказывается почти невозможным по целому ряду причин. Обычно тексты выравнивают по предложениям. Однако при изучении частных лингвистических явлений, в нашем случае, предложных конструкций, требуется выравнивание именно на уровне конструкций.

Актуальность и новизна этого исследования связана не только с разработкой компьютерных методов и средств, но и с изучением морфосинтаксических особенностей русского и венгерского языков, а именно, с точки зрения выражения в них семантико-грамматических предложно-падежных отношений. Для изучения предложно-падежных конструкций русского языка и их эквивалентов в венгерском языке было необходимо создать параллельный русско-венгерский параллельный корпус, что и было сделано.

Была предложена система соответствий русских предлогов и системы падежей в венгерском языке. Для выравнивания русско-венгерских текстов был разработан метод, основанный на теории графов, выявляющий границы предложных конструкций и выделяющий их эквиваленты в венгерских текстах.

К несомненным достоинствам обсуждаемой работы относится ее оригинальность, обработка и анализ большого объема экспериментального материала, разработка программного инструментария. Работа написана хорошим научным языком. Магистрант показал умение работать с литературой, проявил замечательные способности по овладению математическим аппаратом, применяемым в компьютерной лингвистике, по разработке алгоритмов и программ. Разработанные им средства обладают всеми признаками научной новизны и имеют практическое значение.

Работа Кристиана Эгервари выполнена полностью самостоятельно. Полагаю, что ее автор показал высокую научную квалификацию, полностью справился с поставленной задачей, что он является готовым специалистом в области компьютерной лингвистики и заслуживает самой высокой оценки.



«13» июня 2021 г

Захаров В.П.

СОГЛАСИЕ
на обработку персональных данных

Я, Захаров Виктор Павлович,

(фамилия, имя, отчество рецензента)

даю согласие на обработку своих персональных данных оператору - Федеральное государственное бюджетное образовательное учреждение высшего образования «Санкт-Петербургский государственный университет» (далее - СПбГУ), 199034, Санкт-Петербург, Университетская наб., д. 7-9, на следующих условиях:

1. Оператор осуществляет обработку персональных данных исключительно в связи с осуществлением рецензирования и проведения защиты выпускных квалификационных работ обучающихся СПбГУ в целях реализации принципа открытости образовательной деятельности.
1. Перечень персональных данных, передаваемых Оператору на обработку:
 - фамилия, имя, отчество;
 - место работы, должность;
 - ученая степень и звание (при наличии);
 - контактный телефон и адрес электронной почты.
1. Оператор имеет право на обработку персональных данных, то есть совершение, в том числе, следующих действий: обработку (включая сбор, систематизацию, накопление, хранение, уточнение (обновление, изменение), использование, обезличивание, блокирование, уничтожение персональных данных).
2. Данным заявлением разрешаю считать общедоступными, в том числе выставлять в сети Интернет, следующие персональные данные: фамилия, имя, отчество, место работы, должность, ученая степень и звание (при наличии).
3. Обработка персональных данных осуществляется оператором в соответствии с нормами Федерального закона от 27.07.2006 № 152-ФЗ «О персональных данных» и смешанным способом.
4. Срок действия данного Согласия ограничен.

«13» июня 2020 г



Захаров В.П.