

**РЕЦЕНЗИЯ**  
**научного руководителя**  
**на выпускную квалификационную работу обучающегося СПбГУ**  
**Петрушенко Л.С.**  
**по теме «Разработка алгоритма для выравнивания предложных**  
**конструкций в параллельном русско-чешском корпусе»**

Рецензируемая работа Л.С. Петрушенко посвящена выравниванию предложных конструкций в параллельном русско-чешском корпусе. 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. Часто эта задача и не является первоочередной. Обычно тексты выравнивают по предложениям. Однако при изучении частных лингвистических явлений, в нашем случае, предложных конструкций в русском и чешском языках, требуется выравнивание именно на уровне конструкций.

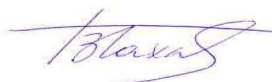
При решении поставленной задачи студентка должна была решить ряд задач, как-то, определить понятие «предложная конструкция», выявить особенности русского и чешского языков и закономерности формирования в них предложных конструкций, создать морфологически размеченный русско-чешский корпус, разработать и запрограммировать алгоритм для выравнивания предложных конструкций.

Все поставленные задачи были решены. В результате в рамках ВКР был создан параллельный корпус с морфоразметкой, составленный из предложений на русском и чешском языках. Были изучены и описаны понятия предлога и предложных конструкций, проведено сравнение русских и чешских первичных предлогов, написана программа выравнивания и проведена оценка качества работы алгоритма. Для решения задач использовались имеющиеся корпуса и инструменты корпусной лингвистики.

Содержание ВКР соответствует заявленной в названии теме. Структура ВКР обоснована и хорошо соотносится с задачами исследования, использованная литература достаточно полно описывает предмет исследования и методы достижения цели. Л.С. Петрушенко показала навыки работы с инструментарием компьютерной лингвистики, умение работать с литературой, владение научным языком и методологией научного анализа.

Не все задуманное получилось сделать, как было задумано. В работе имеются определенные недостатки, например, не проработаны как следует чешские вторичные предлоги в качестве эквивалентов первичных. Это можно объяснить как обилием материала, так и недостаточным знанием чешского.

Работа Л.С. Петрушенко выполнена полностью самостоятельно. Следует отметить большой объем работы по подготовке и проведению экспериментов, по обработке экспериментальных данных. Полагаю, что Л.С. Петрушенко показала высокую научную квалификацию, полностью справилась с поставленной задачей, является обещающим начинающим специалистом в области компьютерной лингвистики и заслуживает высокой оценки за ВКР.



«09» 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.

«09»июня 2020 г



Захаров В.П.