

**Белорусский государственный университет
Факультет социокультурных коммуникаций**

Кафедра компьютерной лингвистики и лингводидактики

**Реферат дипломной работы
«Проблема выравнивания при создании параллельных корпусов
текстов»**

**Гусак Александра Викторовна,
Руководитель Воронович Валерий Викторович**

2015 год

РЕФЕРАТ

Дипломная работа: страниц – 88, в том числе иллюстраций – 4, таблиц – 11, библиографический список – 32 источника, приложений - 3.

Ключевые слова: корпусная лингвистика, корпус текстов, параллельный корпус текстов, автоматическое выравнивание, выравнивание на уровне предложений, ABBYY Aligner, ParaConc.

Объектом дипломной работы является процесс создания параллельных корпусов текстов.

Цель дипломной работы – выявить особенности программного обеспечения, предназначенного для выравнивания параллельных корпусов текстов на уровне предложений.

В ходе выполнения дипломной работы был проведен сравнительный анализ функциональных возможностей программ ABBYY Aligner и ParaConc, а также анализ эффективности их использования для решения проблемы выравнивания параллельных корпусов текстов.

При выполнении работы были использованы следующие **методы:** **анализ** научно-методической литературы и Интернет-источников по проблеме исследования, сопоставительный анализ, метод словарных дефиниций, описательный метод, основанный на анализе программного обеспечения, предназначенного для выравнивания параллельных корпусов текстов.

Автор работы подтверждает, что приведенный в ней расчетно-аналитический материал правильно и объективно отражает состояние исследуемого процесса, а все заимствованные из литературных и других источников теоретические, методологические и методические положения и концепции сопровождаются ссылками на их авторов.

РЭФЕРАТ

Дыпломная работа: старонак – 88, у тым ліку ілюстрацый – 9, табліц – 11, бібліяграфічны слоўнік – 32 крыніцы.

Ключавыя словы: корпусная лінгвістыка, корпус тэкстаў, паралельны корпус тэкстаў, аўтаматычнае выраўноўванне, выраўноўванне на ўзроўні сказаў, ABBYY Aligner, ParaConc.

Аб’ектам дыпломнай работы з’яўляецца працэс стварэння паралельных корпусаў тэкстаў.

Мэта дыпломнай работы – выяўленне асаблівасцяў праграмнага забеспячэння, прызначанага для выраўноўвання паралельных корпусаў тэкстаў.

У ходзе выканання дыпломнай работы быў праведзены параўнальны аналіз функцыйных магчымасцяў праграм ABBYY Aligner і ParaConc, і таксама аналіз эфектыўнасці іх выкарыстання для вырашэння праблемы выраўноўвання паралельных корпусаў тэкстаў.

Пры выкананні работы выкарыстоўваліся наступныя **метады даследавання:** аналіз навукова-метадычнай літаратуры і Інтэрнэт-крыніц па праблеме даследавання, супастаўляльны аналіз, метады слоўнікавых дэфініцый, апісальны метады, заснаваны на аналізе праграмнага забеспячэння, прызначанага для выраўноўвання паралельных карпусоў тэкстаў.

Аўтар работы пацвярджае, што прыведзены разлікова-аналітычны матэрыял дакладна і аб’ектыўна адлюстроўвае становішча працэсу, які даследваецца. А таксама што ўсе запазычаныя з бібліяграфічных крыніц тэарэтычныя, метадалагічныя і метадычныя палажэнні і канцэпцыі суправаджаюцца спасылкамі на іх аўтараў.

ABSTRACT

The diploma paper: pages – 88, including illustrations – 9, tables – 11, references – 32 sources.

Key words: corpus linguistics, corpus, parallel corpus, automatic alignment, sentence alignment, ABBYY Aligner, ParaConc.

The object of the given diploma paper is the process of parallel corpora building.

The purpose of the diploma paper is to indicate the peculiarities of the software aimed at aligning of parallel corpora.

While working on the diploma paper the comparative analysis of the ABBYY Aligner's and ParaConc's functional capabilities was conducted, as well as the analysis of the effective use of the software to solve the problem of parallel corpora alignment.

While working on the diploma paper the following **research methods** were used: analysis of academic literature and Internet-sources, contrastive analysis, method of dictionary definitions, descriptive method, based on the analysis of the software for parallel corpus alignment.

The author of the thesis confirms that the settlement and analytical material given in it correctly and objectively reflects a condition of the studied process, and all theoretical, methodological and methodical provisions and concepts borrowed from literary and other sources are followed by links to their authors.