

Филологические науки

УДК 81'33

Тоирова Гули

доктор философских наук, доцент

кафедра узбекского языкознания

Бухарский государственный университет

Tairova Guli

Doctor of Philosophy, Associate Professor

Department of Uzbek Linguistics

Bukhara State University

ВАЖНОСТЬ ИНТЕРФЕЙСА В СОЗДАНИЕ КОРПУСА IMPORTANCE OF INTERFACE IN CREATING A HOUSING

Аннотация. В статье рассматриваются авторский корпус и его значение в современном глоссарии, мир пушкинского авторского корпуса, чешский писательский корпус, авторский корпус Шекспира и их недостатки. Интерфейс авторского корпуса состоит из различных конструкций и структур, и автор несет ответственность за его полноту, интерфейс должен быть привлекательным и впечатляющим. Создание интерфейса основано на дизайне национальных или современных функций, интерфейс должен включать жизнь и творчество художника в фотографиях. Корпус лингвистики — очень быстро развивающаяся отрасль мира компьютерной лингвистики, которая достигла больших успехов в этом отношении. Корпус лингвистики также преподается как наука в мировых университетах. Предметом этой дисциплины является теория и практика построения корпуса, такие как особенности тела и основы программирования. Корпус лингвистики занимается общей теорией и практикой компьютерной лингвистики, формированием языкового тела и компьютерными технологиями. В статье рассказывается о современных

информационных технологиях, которые создали огромные возможности для языковой функциональности. Компьютерный перевод, редактирование, анализ, электронный словарь и тезаурус являются доказательством нашего мнения. Особенно создание современных электронных словарей и культуры их использования является одним из эффективных способов изучения языка. В частности, роль языковых зданий, созданных и развивающихся быстрыми темпами во всем мире при демонстрации способности и способности владеть языком, очень велика. Цель статьи — изучить лингвистические основы корпуса узбекского языка, изучить лингвистическую ценность лингвистического корпуса, историю корпусного языкознания, изучить авторскую лингвистику корпусов, ее особенности в социальном, лексикологическом, образовательная и другие сферы.

Ключевые слова: *интерфейс, авторский корпус, математическое моделирование, морфологическая и семантическая аннотация, информация, лингвистическая база, искусственный интеллект, компьютерная лингвистика, корпусная лингвистика, языковой корпус, специальное программное обеспечение, электронная библиотека, лексический, морфологический, грамматический, семантические символы, проблемы с лингвистической разметкой.*

Summary. *The article discusses the author's corpus and its significance in the modern glossary, the world of Pushkin's author's corps, the Czech writer's corps, Shakespeare's corpus and their shortcomings. The interface of the author's corpus consists of various designs and structures, and the author is responsible for its completeness, the interface should be attractive and impressive. The creation of the interface is based on the design of national or modern functions, the interface should include the life and work of the artist in photographs. The Linguistics Corps is a very rapidly developing branch of the world of computer linguistics, which has achieved great success in this regard. The Corpus of Linguistics is also taught as a science at world universities. The*

subject of this discipline is the theory and practice of building the corpus, such as body features and programming fundamentals. The Corps of Linguistics deals with the general theory and practice of computer linguistics, the formation of the linguistic body and computer technology. The article talks about modern information technologies that have created tremendous opportunities for language functionality. Computer translation, editing, analysis, electronic dictionary and thesaurus are proof of our opinion. Especially the creation of modern electronic dictionaries and the culture of their use is one of the effective ways to learn the language. In particular, the role of language buildings created and developing rapidly throughout the world in demonstrating the ability and ability to speak the language is very large. The purpose of the article is to study the linguistic foundations of the Uzbek language corpus, to study the linguistic value of the linguistic corpus, the history of corpus linguistics, to study the author linguistics of the corps, its features in social, lexicological, educational and other fields.

Key words: *interface, author's corps, mathematical modeling, morphological and semantic annotation, information, linguistic base, artificial intelligence, computer linguistics, corpus linguistics, language corpus, special software, electronic library, lexical, morphological, grammatical, semantic symbols, problems with linguistic markup.*

В общегосударственной программе большое внимание уделяется как главному направлению общения общеобразовательных школ. В соответствии с этой программой учебные заведения обеспечиваются техникой современных компьютеров. Государственные школы и учебные заведения полностью связаны с Интернетом и Ziyonet. Современные люди овладевают таким количеством информации, что невозможно использовать или обрабатывать информацию без использования информационных и коммуникационных технологий (ИКТ). Год за годом в нашей жизни

развивались компьютерные и информационно-коммуникационные технологии. В настоящее время основная цель образовательной политики направлена на учащихся, важных и необходимых для будущего развития современного образования, отвечающего требованиям общества и власти. Важно привлечь преподавателей и руководителей школ и вузов к развитию профессионального мастерства и с первого дня дополнительного педагогического образования. Существует истина, от которой невозможно отказаться, если представителем настоящего времени, неспособным использовать современные технологии и не способным использовать их для своей жизни, работы и ремесла, считается недостаток. Важно подчеркнуть, что эффективное использование педагогами возможностей современных ИКТ свидетельствует о том, что они являются искусными специалистами.

Тем более, какой-то маленький корпус полностью связан на основе синтаксического анализа. Такое состояние обычно объясняется глубоко или имеет синтаксическую структуру. Например, синтаксический анализ любит большое дерево. Анализ текста в руке занимает много времени. Сейчас он представляет программы на российских и зарубежных сайтах, в которые можно входить прямо и анализировать. Они делятся на независимые и веб-сайты. В последние годы внимание производителей к веб-вложению заслуживает внимания. У этих систем есть некоторые преимущества: в то же время несколько пользователей могут анализировать один и тот же документ, он не требует установки дополнительных программ, но кроме него установлен браузер, право входа ограничено, можно наблюдать за процессом анализа. Словоформа, лемма и тег принадлежат к морфологической анализирующей системе. Словоформа — это морфологическая единица выбранного текста. Первый шаг анализа словоформы — это лемминг или придание лексемой формы слова. Самым сложным этапом настройки языков флексии является лемминг или лексема в форме слова для соединения с формой слова в виде тега. Потому что в

инфлексивных языках грамматическое значение слова формы смешивается с корнем слова. В отличие от флективных языков лемминг гораздо проще в агглютинативном языке [1, с. 95–102.]. Часть без грамматической формы слова равна лемме о корне или базовом фундаменте. В настройках лемма указана внутри следующего знака:<*>. Во всех частях речи лемминг выглядит следующим образом или, если он основан на тенденции «часть основания слова равна лемме», в группе глаголов глагол лемма дается как форма императивного настроения II человека. В статьях словаря они даны как инфинитив. Но он не подходит для корпусов, потому что в тексте корпусов ищется не форма, а форма. Согласно этому глаголу лемма имеет следующую форму: учить<читать>, не быть<быть>, показывать<видеть> [5, с. 49–50]. Во время настройки требуется написать от 5 до 10 морфологических тегов (объяснений), иногда больше, чем их для каждой словоформы.

Корпус состоит из набора электронных текстов, работающих на специальной поисковой системе и миллионов контекстов ключевых слов. Авторский корпус основан на текстах, содержащих художественный, публицистический и эпический жанры, созданные конкретным художником. Это серия лексикографических источников, которые включают в себя различные словари (основной, частотный, топонимика, грамматика, фразы и т. Д.). Авторский корпус имеет большое значение в современном дизайне. Этот корпус является ресурсом для создания больших словарей. Со временем корпус стал мощным информационным ресурсом, который важен для различных языковых тенденций. Из-за компьютерных трупов словари создаются и обрабатываются быстрее по сравнению с предыдущим. Таким образом, словарь отражает язык от начала до конца словаря (до публикации), и в словаре нет «устаревшего» слова. Корпус должен иметь уникальный интерфейс [9].

Что такое интерфейс? Термин «интерфейс» происходит от английского языка и используется для обозначения «внешнего вида». Это слово часто используется в компьютерных технологиях. Это единственная система связи, которая обеспечивает разнообразные коммуникации между человеком и автомобилем. Интерфейс представляет собой систему связи между различными блоками и сложными блоками, а также технологией и пользователем. Это логическое (система представления информации) и физическое (передача данных). Интерьеры авторского корпуса имеют разный дизайн и структуру, а их совершенствование поручено создателю. Потому что интерфейс — это первое впечатление о теле. При создании интерфейса необходимо учитывать национальные декоративные узоры и характеристики, отражающие классику или современность. Жизнь и деятельность создателя, созданные им работы в отдельном окне, а также фотогалерея должны быть на интерфейсе [8].

Вышеуказанные материалы позволяют использовать тело в качестве электронной библиотеки. В системе «Поиск» жизнь создателя, деятельность, работы, созданные им в отдельном окне, и окна, такие как фотографии, используют возможности тела. Эти окна являются программным обеспечением, которое ищет корпуса. Корпус автора должен быть дружелюбным, оптимально универсальным и совершенным. Глобальная сеть характеризуется наличием больших объемов материалов (множество материалов) и глубиной его. Например, корпуса Пушкина и Чехова являются идеальными с морфологическими и семантическими аннотациями. Но это не отвечает современным требованиям программирования. Потому что им было 8 или 10 лет. Кроме того, превосходный дизайн корпуса Шекспира отличается дизайном всего дизайна, но он не является морфологическим или семантическим. Каждый из существующих трупов имеет явное меньшинство. Например, таджикский язык может функционировать только как электронная библиотека, фото и

видео галерея [2, с.231-234] Это корпус совершенно не имеет себе равных, а промежуточная система несовершенна и неполна. Но это может указывать только на один кусок работы. Пушкинский и чеховский труп, созданные русскими лингвистами, не имеют синтаксического значения. Труп автора — это электронная база данных Абдуллы Каххора «История пациента», основанная на искусственном интеллекте: обработка узбекского языка посредством текстового редактирования, автоматического редактирования и перевода, поиска различных символов. Созданный орган — может определять период и частоту слова, анализировать живую конструкцию языка, возможность комбинировать слова, автоматически обрабатывать текст (например, переводить) и автоматически выполнять различные поисковые системы. Это экономит время. Примечание. В настоящее время пользователи используют простой карт ридер, а не базу данных linux. Имеет следующие интерфейсы: Т интерфейс играет ключевую роль в создании трупа. Необходимо создать простую в использовании рабочую среду, отвечающую требованиям современного дизайна программного обеспечения.

В заключение, лингвистическое моделирование тегов целесообразно, потому что в лингвистическом моделировании морфологический тег превращается в обычную форму аббревиатуры. Вырабатывается формы специальной лингвистической модели для настройки каждой части речи. Научная новизна исследования заключается в следующем:

- 1) впервые в узбекской лингвистике раскрыт термин корпус, особенности корпуса, теоретические основы, лингвистическое, практическое и образовательное значение языкового корпуса;
- 2) становление и развитие корпусной лингвистики, особенности первого и последующих поколений, история языковых трупов на русском и английском языках, состояние корпусной лингвистики, особенности современного русского, английского, турецкого и таджикского

языков, их общий и разнообразный аспекты, классификация типов корпусов;

- 3) общие принципы построения тела: технологический процесс проектирования и этапы построения тела, важность формирования тела, тот факт, что метка является лингвистическим инструментом, характеризующимся характерными чертами менеджера корпуса и его типов.

Литература

1. Ванюшкин А.С., Гращенко Л.А. Оценка алгоритмов извлечения ключевых слов: инструментарий и ресурсы // Новые информационные технологии в автоматизированных системах. 2017. № 20. С. 95–102.
2. Николаев И.С., Митренина О.В., Ландо Т.М. Прикладная и компьютерная лингвистика. М. URSS, 2016. 320 с.
3. Недошивина Е.В. Программы для работы с корпусами текстов: обзор основных корпусных менеджеров. Учебно-методическое пособие. Санкт-Петербург. 2006. 26 с.
4. Namroyeva Sh. Linguistic basis for the creation of the Uzbek language. 2018. 52 б.
5. Namroyeva Sh. Use in education from the corpus “Language and literary education” Journal. September 2017. № 9. PP. 49-50.
6. Мамонтова В. В. Корпус параллельных текстов и база данных для исследования переводческих соответствий: проблемы и процедуры формирования. URL: www.gramota.net/materials/1/2008/8-2/50.html
7. Antonova A., Alexey M. Building a Web-based parallel corpus and filtering out machinetranslated text. URL: <https://www.aclweb.org/anthology/W11-1218>
8. Toirova G. Actual problems of Uzbek linguistic research. International Scientific Journal Theoretical & Applied Science p-ISSN: 2308-4944

(print) e-ISSN: 2409-0085 (online) Year: 2019 Issue: 07 Volume: 75
Published: 30.07.2019 Impact Factor 6,630.

9. Toirova G. Importance of Interface in Creating Corpus. International Journal of Recent Technology and Engineering (IJRTE) ISSN: 2277-3878, Volume-8 Issue-2S10, September 2019.
10. Toirova G. The Role Of Setting In Linguistic Modeling. International Multilingual Journal of Science and Technology. ISSN: 2528-9810 Vol. 4 Issue 9, September – 2019. PP. 722-723.
11. Рыков В.В. Курс лекций по корпусной лингвистике. URL: <http://rykov-cl.narod.ru/c.html>.