

Создание авторских корпусов: "Корпус текстов Имадеддина Насими"

Халилова Гамида Азизага гызы

Сотрудник

Бакинский филиал Московского государственного университета имени М.В.Ломоносова,
Баку, Азербайджан

E-mail: hamida.khalilova@mail.ru

Создание авторских корпусов носит новаторский характер. В России, к примеру, созданы корпуса текстов Пушкина, Чехова, в Англии – корпус текстов Шекспира, в Азербайджане – Поэтический корпус текстов Физули [Халилова 2024], Корпус текстов Низами [Халилова 2023], Корпус текстов М.А. Сабира [Халилова 2023]. Авторские корпуса особенно активно создаются в последнее время, о них также писал в своем труде В.П. Захаров [Захаров 2005]. Авторские корпуса позволяют исследовать особенности отдельной языковой личности. В текстах одной личности язык в полной мере реализуется как целостная система. Авторские корпуса могут быть снабжены словарями, через которые можно находить нужные контексты.

Нужно также отметить, что благодаря корпусу текстов исследователь может работать с массивом текстов, имея возможность извлечь оттуда интересующую его лексическую единицу, то есть может проанализировать, выявить поведение данной единицы в естественной языковой среде, не теряя при этом время на сбор материала, разметку и т. д. [Халилова 2024].

Актуальность и новизна исследования обусловлены значимостью создания корпусов текстов азербайджанского языка, который на данный момент мало представлен корпусами текстов. Следует особо подчеркнуть, что создание корпусов текстов отдельных азербайджанских авторов, корпуса художественных текстов не менее важно и для создания Национального корпуса (НК) азербайджанского языка, создание которого представляет собой стратегический вопрос государственного значения. Итак, НК включают в себя все стили и жанры. В НК входят и произведения художественной литературы, так как именно произведения художественной литературы и устного народного творчества имеют культурную значимость, в них также заключены прецедентные феномены [Халилова 2025].

В значительной степени влияет на актуальность данной работы и выбор материала – художественные тексты Имадеддина Насими, написанные на азербайджанском языке. При этом данные тексты ранее не являлись объектом корпусного и лингвостатистического исследования. Одним из мыслителей, сыгравших исключительную роль в развитии азербайджанской поэзии на родном языке, является именно Имадеддин Насими. Он заложил основы философской поэзии в истории азербайджанской литературы, обогатил художественное слово как по форме, так и по содержанию и поднял азербайджанский язык до уровня литературно-художественного языка.

Результаты исследования:

В Корпус текстов Имадеддина Насими включены газели, общественно-философские стихотворения, мустазады, маснави и другие поэтические образцы, написанные им на родном азербайджанском языке и представленные в двухтомном издании. В корпус текстов И. Насими входят: газели – 287, общественные и философские стихотворения – 155, мустазады – 4, алифлам и обратный алфавит – 5, тарджибенды – 3, маснави – 2, мулемма – 1, туюги – 1, дополнения (дополнительные произведения) – 15.

В результате исследования: определено общее количество токенов (word tokens), то есть употреблений слов, и типов слов (word types), то есть словоформ; составлен частотный

список всех употреблений слов; составлен конкорданс всех употреблений слов; определены словоформы с наибольшей и наименьшей частотностью, выявлены единицы, представляющие собой индивидуально-авторское употребление.

В рамках данного исследования также начата работа по созданию параллельного корпуса, включающего оригинальные тексты Имадеддина Насими на азербайджанском языке и их переводы на русский язык. Переведённая литература и многоязычные параллельные корпуса являются незаменимым ресурсом для изучения международных связей азербайджанской литературы и процессов перевода. С помощью искусственного интеллекта в таких корпусах становится возможным автоматизированный анализ таких вопросов, как переводческие стратегии, стилистическое соответствие и адаптация культурных элементов.

Таким образом, данная работа является продолжением исследований по созданию корпусов текстов азербайджанского языка. Создание корпуса текстов Имадеддина Насими представляет собой важный этап в развитии корпусных исследований азербайджанского языка. Разработанный корпус позволит проводить системный анализ лексических, морфологических и синтаксических особенностей поэтического языка автора, а также применять современные методы компьютерной обработки текста. Результаты данной работы могут быть использованы в дальнейших лексикографических и корпусных исследованиях, а также служить основой для новых научных разработок в области языкознания, перевода и цифровых гуманитарных исследований.

Источники и литература

- 1) Захаров В.П. Корпусная лингвистика: Учебно-метод. пособие. – СПб., 2005. – 48 с.
- 2) Халилова Г.А. Корпусное исследование на материале азербайджанского и русского языков. Создание корпуса текстов Низами Гянджеви // – М.: Текст. Исследования молодых учёных, – 2023. № 1–2 (8–9), с. 121-128.
- 3) Халилова Г.А. Корпусное исследование сатир Мирза Алекпера Сабира. Создание корпуса текстов Сабира // *Filologiya məsələləri*, – Bakı: Elm v təhsil, – 2023. №14, – s. 425-432.
- 4) Халилова Г.А. Создание поэтического корпуса газелей М. Физули // – М.: Балтийский гуманитарный журнал, – 2024. Т. 13. № 1(46), – с. 113-116.
- 5) Халилова Г.А. Художественный текст как объект лингвостатистического исследования автореф. дис. ... канд. филол. наук. Баку: Национальная Академия наук Азербайджана, 2025, 37 с.