

МРНТИ 16.31.02
УДК 334:005.342

<https://doi.org/10.46914/2959-3999-2025-1-1-45-50>

ЖАППАР К.З.,*¹

к.ф.н., ассоциированный профессор.

*e-mail: zhappar.karlygash@mail.ru

ORCID ID: 0000-0001-8655-7675

КУРМАНБАЕВА А.А.,¹

к.ф.н., ассоциированный профессор.

e-mail: al.kurmanbaeva@turan-edu.kz

ORCID ID: 0000-0001-5174-5086

ТОМАНОВА Н.М.,¹

к.ф.н., ассоциированный профессор.

e-mail: n.tomanova@turan-edu.kz

ORCID ID: 0009-0004-1730-6150

АЛДАБЕРГЕНКЫЗЫ Л.,¹

м.ф.н., сениор-лектор.

e-mail: l.aldabergenkyzy@turan-edu.kz

ORCID ID: 0000-0003-3895-2085

¹Университет «Туран»,

г. Алматы, Казахстан

НАЦИОНАЛЬНЫЙ КОРПУС КАЗАХСКОГО И РУССКОГО ЯЗЫКОВ: ПЕДАГОГИЧЕСКИЕ АСПЕКТЫ ПРИМЕНЕНИЯ

Аннотация

В данной статье рассматриваются педагогические аспекты применения Национальных корпусов русского и казахского языков. Проведен сравнительный анализ их структуры, текстовой базы, методов аннотации и возможностей использования в образовательном процессе. Результаты исследования подтверждают важность корпусов для изучения языка, грамматического и лексического анализа, а также формирования переводческих навыков. Установлено, что Национальный корпус казахского языка находится в стадии развития и требует дальнейшего совершенствования для полноценной интеграции в образовательную среду. Также обсуждается необходимость разработки методических рекомендаций по эффективному использованию корпусной лингвистики в преподавании.

Ключевые слова: национальный корпус, корпусная лингвистика, корпус казахского языка, корпус русского языка, образование, текстовый анализ, перевод, изучение языка, цифровые технологии.

Введение

Современные технологии значительно изменили подходы к изучению и преподаванию языков. Корпусная лингвистика, как одна из перспективных областей исследования, предоставляет преподавателям и студентам мощные инструменты для анализа языкового материала. Национальные корпуса русского и казахского языков, представляющие собой обширные базы текстов с различными видами разметки, открывают новые возможности для лингвистического образования, методической подготовки педагогов и научных исследований.

Национальный корпус русского языка (НКРЯ) активно используется в образовательном процессе благодаря своим расширенным поисковым возможностям, разнообразию текстов и наличию аннотированных примеров. В то же время Национальный корпус казахского языка (НККЯ), находящийся в стадии развития, требует дополнительных улучшений в области методических рекомендаций для педагогического применения.

Актуальность исследования обусловлена необходимостью интеграции национальных корпусов в образовательный процесс, разработки эффективных методик их использования в

преподавании русского и казахского языков, а также формирования цифровых компетенций у студентов и преподавателей. В статье рассматриваются основные педагогические аспекты применения НКРЯ и НККЯ, анализируются их возможности для изучения грамматики, лексики, стилистики и перевода, а также предлагаются рекомендации по их эффективному использованию в учебном процессе.

Материалы и методы

В исследовании использовались данные Национального корпуса русского языка (НКРЯ) и Национального корпуса казахского языка (НККЯ). Для анализа педагогических аспектов их применения рассмотрены:

- ♦ структура и функциональные возможности корпусов, включая объем текстов, наличие аннотации, возможности поиска и уровень автоматизированной обработки данных;
- ♦ методы интеграции корпусов в образовательный процесс, основанные на существующих исследованиях в области корпусной лингвистики и педагогики;
- ♦ опыт использования корпусных данных в преподавании, изученный на основе научных публикаций и практических методик преподавателей.

В работе применены следующие методы исследования: контент-анализ – изучение научных публикаций по корпусной лингвистике и методике преподавания языков; сравнительный анализ – сопоставление функциональных возможностей НКРЯ и НККЯ, а также их применимости в образовательной деятельности; метод кейс-стади – анализ примеров использования корпусов в учебном процессе, включая преподавание русского языка как иностранного и изучение казахского языка; экспертное интервью – сбор мнений преподавателей и специалистов в области корпусной лингвистики о перспективах применения корпусов в обучении.

Исследование направлено на выявление эффективных методов использования корпусных технологий в педагогической практике, а также на разработку рекомендаций по внедрению корпусных ресурсов в учебные программы.

Результаты и обсуждение

Рассмотрим возможности Национального корпуса русского и казахского языков в образовательной деятельности.

На основе сравнительного анализа было выявлено, что Национальный корпус русского языка (НКРЯ) обладает развитой структурой и широкими возможностями для образовательного применения. Он предоставляет доступ к различным видам разметки (морфологической, синтаксической, семантической), поддерживает сложные поисковые запросы и содержит большой объем текстов разных жанров. Это делает его полезным инструментом для преподавания грамматики, стилистики, лексикологии и даже перевода.

В отличие от него, Национальный корпус казахского языка (НККЯ) пока имеет ограниченный объем текстов и менее развитую систему разметки [1]. Основные трудности связаны с агглютинативной природой казахского языка, что затрудняет автоматизированный морфологический анализ. Однако потенциал НККЯ в образовательном процессе значителен, особенно в контексте изучения казахского языка как второго языка, развития переводческих навыков и анализа особенностей письменной речи [2].

Таблица 1 – Сравнительная характеристика образовательных возможностей НКРЯ и НККЯ

Параметр	Национальный корпус русского языка (НКРЯ)	Национальный корпус казахского языка (НККЯ)
Объем текстов	1,2 млрд словоформ	Около 20 млн словоформ
Жанровое разнообразие	Художественная, научная, публицистическая, деловая, устная речь и др.	Преимущественно художественная и официально-деловая речь
Морфологическая разметка	Полноценная аннотация	Частичная аннотация, ошибки в обработке агглютинативных форм

Продолжение таблицы 1

Синтаксическая разметка	Поддерживается	Не реализована
Возможность поиска по корпусу	Расширенные функции поиска, включая лемматизацию и контекстный анализ	Поиск ограничен, без сложных запросов
Использование в образовательных целях	Широкое применение в преподавании русского языка и лингвистических исследованиях	Ограниченное применение, потенциал для изучения казахского языка
Примечание: Составлено авторами.		

Вопрос о применении национальных корпусов в образовательной сфере нашей страны становится все более актуальным в современной педагогической науке [5].

На основе проведенного исследования выявлены основные способы использования национальных корпусов в преподавании:

- ♦ Разбор грамматических конструкций: преподаватели могут использовать корпуса для поиска примеров употребления сложных грамматических явлений. В НКРЯ доступны примеры с аннотацией, что облегчает работу [3]. В НККЯ эта функция пока ограничена, но может развиваться.

- ♦ Изучение лексических единиц и стилистики: поиск по корпусу позволяет анализировать употребление слов в различных контекстах. Это особенно полезно для изучающих казахский язык, так как помогает увидеть живые примеры использования слов и выражений.

- ♦ Работа с аутентичными текстами: студенты могут использовать корпусные данные для анализа текстов в оригинале, что способствует развитию навыков чтения и критического анализа.

- ♦ Переводческая практика: параллельные тексты в корпусах позволяют анализировать особенности перевода и выявлять типичные ошибки при передаче смыслов между русским и казахским языками.

- ♦ Автоматизированные инструменты анализа: использование корпусных технологий в цифровом обучении может повысить эффективность самостоятельной работы студентов и облегчить создание учебных материалов.

Наряду с названными возможностями и значительным потенциалом, существует ряд проблем на уровне развития перспектив внедрения корпусных данных в образование, в их числе:

- ♦ ограниченные технические возможности НККЯ по сравнению с НКРЯ;
- ♦ недостаточная информированность преподавателей о методах использования корпусных данных в обучении;
- ♦ необходимость разработки учебных пособий и методических рекомендаций, основанных на корпусных исследованиях.

Для дальнейшего развития области применения национальных корпусов казахского и русского языков в сфере образования республики могут быть предложены следующие направления: расширение текстовой базы НККЯ, включение текстов разговорной речи и научной литературы; автоматизация морфологической и синтаксической разметки казахского языка с использованием методов машинного обучения; разработка методических материалов и тренингов для преподавателей по использованию корпусных данных в педагогической практике.

Таким образом, национальные корпуса русского и казахского языков представляют собой мощные инструменты для обучения, однако требуют дальнейшего развития и адаптации к образовательным потребностям.

Заключение

Рассмотрение педагогических аспектов применения Национального корпуса русского языка (НКРЯ) и Национального корпуса казахского языка (НККЯ) показало, что корпусные технологии могут значительно повысить качество языкового образования. Они позволяют проводить

детальный анализ грамматических и лексических структур, способствуют развитию переводческих навыков и помогают обучающимся работать с аутентичными текстами.

Национальный корпус русского языка обладает широкими возможностями для образовательного процесса благодаря большому объему текстов, развитой системе разметки и продвинутым инструментам поиска. В то же время Национальный корпус казахского языка находится в стадии активного развития и пока уступает НКРЯ по ряду параметров, включая объем текстовой базы, уровень аннотации и поисковые функции. Однако он имеет значительный потенциал, особенно в изучении казахского языка как второго языка и развитии цифровых инструментов для автоматизированного анализа текста.

Несмотря на очевидные преимущества, интеграция национальных корпусов в образовательный процесс требует решения ряда проблем:

- ♦ доработки НККЯ с учетом морфологических и синтаксических особенностей казахского языка;
- ♦ разработки методических рекомендаций по использованию корпусов в обучении;
- ♦ повышения цифровой компетентности преподавателей в области корпусной лингвистики.

Перспективы дальнейших исследований включают разработку автоматизированных инструментов для анализа текстов казахского языка, расширение текстовой базы НККЯ и внедрение корпусных методов в цифровые образовательные платформы. Внедрение корпусных технологий в преподавание русского и казахского языков будет способствовать повышению качества обучения, развитию академических компетенций обучающихся и созданию новых подходов к анализу языкового материала.

ЛИТЕРАТУРА

- 1 Фазылжан А. Концептуальные основы Национального корпуса казахского языка // Этносаясат. – 2023. – № 2. – С. 7–15. URL: <https://journal.etnosayasat.kz/index.php/etp/article/view/29>
- 2 Искакова А., Курмангалиева А., Сулейменова Э., Искакова Г. Алматинский корпус казахского языка: перспективы и результаты. – Алматы: Ulagat, 2020. URL: <https://ulagat.com/2020/12/28/алматинский-корпус-казахского-языка/>
- 3 Соколова Е.А. Использование Национального корпуса русского языка при формировании переводческих компетенций // Гуманитарный научный вестник науки. – 2021. – № 89–94. URL: <https://naukavestnik.ru/doc/2021/08/Sokolova.pdf>
- 4 Шаповалов М.И. Национальный корпус русского языка: как использовать сервис в школе? // URL: <https://ruscorpora.ru/media/publications/10e81e693567d150178ab38c63c2fa01.pdf>
- 5 Мурсал А. Эффективные способы использования национального корпуса казахского языка в обучении. – Алматы: Казахский национальный педагогический университет имени Абая, 2024. URL: <https://www.kaznpu.kz/ru/36850/news/>
- 6 Бейсенова А. К вопросу обучения русскому языку в Казахстане // Профессиональное образование в современном мире. – 2023. – № 4. URL: <https://www.ppfsvfu.ru/jour/article/view/289>
- 7 Жұбанов А. Қазақ тілінің ұлттық корпусы және метабелгіленім мәселесі // Тілтаным. – 2015. – № 1. – 21–29 бб. URL: <https://www.tiltanym.kz/jour/article/view/312/145>

REFERENCES

- 1 Fazylzhan A. (2023) Konceptual'nye osnovy Nacional'nogo korpusa kazahskogo jazyka // Jetnosajasat. No. 2. P. 7–15. URL: <https://journal.etnosayasat.kz/index.php/etp/article/view/29>. (In Russian).
- 2 Iskakova A., Kurmangalieva A., Sulejmenova Je., Iskakova G. (2020) Almatinskij korpus kazahskogo jazyka: perspektivy i rezul'taty. Almaty: Ulagat. URL: <https://ulagat.com/2020/12/28/almatinskij-korpus-kazahskogo-jazyka/>. (In Russian).
- 3 Sokolova E.A. (2021) Ispol'zovanie Nacional'nogo korpusa russkogo jazyka pri formirovanii perevodcheskih kompetencij // Gumanitarnyj nauchnyj vestnik nauki. No. 89–94. URL: <https://naukavestnik.ru/doc/2021/08/Sokolova.pdf>. (In Russian).

4 Shapovalov M.I. Nacional'nyj korpus ruskogo jazyka: kak ispol'zovat' servis v shkole? // URL: <https://ruscorpora.ru/media/publications/10e81e693567d150178ab38c63c2fa01.pdf>. (In Russian).

5 Mursal A. (2024) Jeffektivnye sposoby ispol'zovanija nacional'nogo korpusa kazahskogo jazyka v obuchenii. Almaty: Kazahskij nacional'nyj pedagogicheskij universitet imeni Abaja. URL: <https://www.kaznpu.kz/ru/36850/news/>. (In Russian).

6 Bejsenova A. (2023) K voprosu obuchenija russkomu jazyku v Kazahstane // Professional'noe obrazovanie v sovremennom mire. No. 4. URL: <https://www.ppfsvfu.ru/jour/article/view/289>. (In Russian).

7 Jūbanov A. (2015) Qazaq tılınıń ultiq korpusy jäne metabelgilenim мәсесі // Tiltanym. No. 1. 21–29 pp. URL: <https://www.tiltanym.kz/jour/article/view/312/145>. (In Kazakh).

ЖАППАР Қ.З.,*¹

ф.ғ.к., қауымдастырылған профессор.

*e-mail: zhappar.karlygash@mail.ru

ORCID ID: 0000-0001-8655-7675

ҚҰРМАНБАЕВА А.А.,¹

ф.ғ.к., қауымдастырылған профессор.

e-mail: al.kurmanbaeva@turan-edu.kz

ORCID ID: 0000-0001-5174-5086

ТОМАНОВА Н.М.,¹

ф.ғ.к., қауымдастырылған профессор.

e-mail: n.tomanova@turan-edu.kz

ORCID ID: 0009-0004-1730-6150

АЛДАБЕРГЕНҚЫЗЫ Л.,¹

ф.ғ.м., сениор-лектор.

e-mail: l.aldabergenkyzy@turan-edu.kz

ORCID ID: 0000-0003-3895-2085

¹ Университет «Туран»,

г. Алматы, Қазақстан

ҚАЗАҚ ЖӘНЕ ОРЫС ТІЛДЕРІНІҢ ҰЛТТЫҚ КОРПУСЫ: ПЕДАГОГИКАЛЫҚ ҚОЛДАНУ АСПЕКТІЛЕРІ

Аңдатпа

Бұл мақалада орыс және қазақ тілдерінің ұлттық корпустарын педагогикалық тұрғыдан қолдану мәселесі қарастырылады. Ұлттық корпустардың құрылымы, мәтіндік базасы, аннотация әдістері және білім беру үдерісінде қолдану мүмкіндіктері салыстырмалы түрде талданады. Зерттеу нәтижелері ұлттық корпустардың тіл үйретуде, грамматикалық және лексикалық талдауда, аударма дағдыларын қалыптастыруда маңыздылығын көрсетеді. Қазақ тілінің корпусының әлі де даму сатысында тұрғандығы және оның оқу процесіне толық интеграциялануы үшін қажетті жетілдірулер ұсынылады. Сонымен қатар, білім беру саласында корпус лингвистикасын тиімді пайдалану үшін әдістемелік нұсқаулар жасау қажеттілігі талқыланады.

Тірек сөздер: ұлттық корпус, корпус лингвистикасы, қазақ тілі корпусы, орыс тілі корпусы, білім беру, мәтіндік талдау, аударма, тіл үйрету, цифрлық технологиялар.

ZHAPPAR K.Z.,*¹

c.ph.s., associate professor.

*e-mail: zhappar.karlygash@mail.ru

ORCID ID: 0000-0001-8655-7675

KURMANBAEVA A.A.,¹

c.ph.s., associate professor.

e-mail: al.kurmanbaeva@turan-edu.kz

ORCID ID: 0000-0001-5174-5086

TOMANOVA N.M.,¹

c.ph.s., associate professor.

e-mail: n.tomanova@turan-edu.kz

ORCID ID: 0009-0004-1730-6150

ALDABERGENKYZY L.,¹

m.ph.s., senior lecturer.

e-mail: l.aldabergenkyzy@turan-edu.kz

ORCID ID: 0000-0003-3895-2085

¹Turan University,

Almaty, Kazakhstan

THE NATIONAL CORPUS OF THE KAZAKH AND RUSSIAN LANGUAGES: PEDAGOGICAL ASPECTS OF APPLICATION

Abstract

This article explores the pedagogical aspects of using the National Corpora of the Russian and Kazakh languages. It provides a comparative analysis of their structure, text database, annotation methods, and potential applications in education. The study highlights the importance of corpora in language learning, grammatical and lexical analysis, and the development of translation skills. The findings indicate that the Kazakh language corpus is still in development and requires further improvements for full integration into the educational process. Additionally, the need for methodological guidelines on effectively utilizing corpus linguistics in education is discussed.

Key words: national corpus, corpus linguistics, Kazakh language corpus, Russian language corpus, education, text analysis, translation, language learning, digital technologies.