

A.A. Baimyrza 

A. Baitursynuly Institute of Linguistics, Almaty, Kazakhstan
e-mail: baimurzina.ainur@gmail.com

DIGITAL LINGUODIDACTICS AND NATIONAL IDENTITY: INTEGRATING THE ABAI POETIC CORPUS INTO LANGUAGE AND LITERATURE EDUCATION

This article explores the potential of the Abai Poetic Corpus, an open-access resource, as an effective tool for promoting national values in the modern educational process. The study consisted of two main stages: the first involved linguistic analysis based on the corpus, and the second was a pedagogical experiment conducted among senior secondary school pupils and first-year university students. At the linguistic stage, AntConc software was used to perform frequency, collocation, and concordance analyses. Several key concepts, systematically found in the works of Abai Kunanbayev, were identified: adam (a universal humanist category), aqyl (a measure of intellect and wisdom), iman (spiritual faith and moral purity), and namys (personal honor and social responsibility). These notions were shown to form the foundation of Kazakh ethical thought. At the pedagogical stage, students completed specially designed corpus-based tasks. During a three-lesson pilot experiment, it was observed that learners developed a deeper understanding of Abai's moral triad – aqyl, qairat, zhurek (intellect, willpower, and heart) – and formed a renewed perspective on cultural heritage. Both quantitative and qualitative methods were applied, ensuring a comprehensive research approach. Furthermore, corpus-based methods helped enhance students' skills in independent text work, data comparison, and critical thinking. The findings indicate that digital linguistic resources provide real opportunities for revitalizing national identity and enriching educational content. The results also contribute valuable insights to the field of digital linguodidactics.

Keywords: Abai Kunanbayev, corpus linguistics, poetic corpus, national identity, moral education, digital pedagogy.

А.А. Баймырза

А. Байтұрсынұлы атындағы Тіл білімі институты, Алматы, Қазақстан
e-mail: baimurzina.ainur@gmail.com

Цифрлық лингводидактика және ұлттық бірегейлік: Абай поэтикалық корпусын тіл мен әдебиет пәндеріне интеграциялау

Бұл мақалада қазіргі заманғы білім беру үдерісінде ұлттық құндылықтарды дәріптеудің тиімді жолы ретінде ашық қолжетімді Абай поэзия корпусының қолданылу мүмкіндіктері қарастырылады. Зерттеу жұмысы екі негізгі кезеңнен тұрды: біріншісі – корпус материалдарына сүйенген лингвистикалық талдау, екіншісі – орта мектептің жоғары сынып оқушылары мен университеттің бірінші курс студенттеріне жүргізілген педагогикалық тәжірибе. Лингвистикалық кезеңде AntConc бағдарламасы пайдаланып, сөз жиілігі, колокация және конкорданс талдаулары жасалды. Абай Құнанбайұлы шығармаларында жүйелі түрде кездесетін негізгі ұғымдар анықталды: адам (жалпыадамзаттық гуманистік категория), ақыл (зерде мен парасаттылық өлшемі), иман (рухани сенім мен ар тазалығы), намыс (жеке тұлғаның ар-намысы, қоғам алдындағы жауапкершілік). Бұл ұғымдар қазақ этикалық ойлау жүйесінің іргетасы екені дәлелденді. Педагогикалық кезеңде студенттер корпус негізінде арнайы әзірленген тапсырмаларды орындады. Үш сабақтан тұратын шағын тәжірибе барысында олардың Абайдың моральдық триадасын – ақыл, қайрат, жүрек – тереңірек түсініп, мәдени мұраға жаңаша көзқарас қалыптастырғаны байқалды. Сандық және сапалық әдістер қатар қолданылып, кешенді зерттеу жүргізілді. Сонымен қатар, корпустық әдістер оқушылардың мәтінмен өздігінен жұмыс істеу, деректерді салыстыру, сыни тұрғыдан ойлау дағдыларын дамытуға ықпал етті. Қорытындысында, цифрлық лингвистикалық ресурстар ұлттық болмысты жаңғыртуға және білім беру мазмұнын байытуға нақты мүмкіндік беретінін көрсетті. Алынған зерттеу нәтижелері цифрлық лингводидактика үшін практикалық құнды мәліметтер қатарын толықтырады.

Түйін сөздер: Абай Құнанбайұлы, корпустық лингвистика, поэтикалық корпус, ұлттық бірегейлік, адамгершілік тәрбиесі, цифрлық педагогика.

А.А. Баймырза

Институт языкознания им. А. Байтурсынулы, Алматы, Казахстан
e-mail: baimurzina.ainur@gmail.com

Цифровая лингводидактика и национальная идентичность: интеграция поэтического корпуса Абая в обучение языку и литературе

В данной статье рассматриваются возможности использования открытого Поэтического корпуса Абая как эффективного инструмента популяризации национальных ценностей в современном образовательном процессе. Исследование состояло из двух основных этапов: первый – лингвистический анализ на материале корпуса, второй – педагогический эксперимент, проведённый среди учащихся старших классов и студентов первого курса университета. На лингвистическом этапе с применением программы AntConc были выполнены анализ частотности слов, колокаций и конкордансов. В произведениях Абая Кунанбаева были выявлены ключевые понятия, встречающиеся систематически: адам (общечеловеческая гуманистическая категория), ақыл (мерило разума и мудрости), иман (духовная вера и нравственная чистота), намыс (честь личности и ответственность перед обществом). Доказано, что данные категории являются фундаментом казахской этической мысли. На педагогическом этапе студенты выполняли специально разработанные задания на основе корпуса. В ходе трёхурочного мини-эксперимента было установлено, что они глубже осознали моральную триаду Абая – ақыл, қайрат, жүрек – и сформировали новое отношение к культурному наследию. В исследовании сочетались количественные и качественные методы, что обеспечило комплексный подход. Кроме того, использование корпусных технологий способствовало развитию у обучающихся навыков самостоятельной работы с текстом, сопоставления данных и критического мышления. В заключение делается вывод, что цифровые лингвистические ресурсы открывают реальные возможности для обновления национальной идентичности и обогащения содержания образования. Полученные результаты дополняют ряд практических данных для цифровой лингводидактики.

Ключевые слова: Абай Кунанбаев, корпусная лингвистика, поэтический корпус, национальная идентичность, нравственное воспитание, цифровая педагогика.

Introduction

In today's digital age, the way language and literature are taught is changing rapidly. Instead of relying only on textbooks and teacher-centered lessons, many classrooms are moving toward interactive, student-focused, and technology-supported methods. This change is happening not only because new digital tools are available, but also because educators see the need for more relevant, inclusive, and culturally sensitive teaching. One of the new directions in this area is called *digital linguodidactics*. It brings together digital corpora, annotation tools, and computational approaches to help students explore language, literature, and culture in new ways. These tools allow learners to work with authentic texts, build independence, and engage more deeply with meaning.

In Kazakhstan, these changes have special importance. Language education here is not only about learning grammar and vocabulary, but also about passing on cultural memory, moral values, and a sense of national identity. The legacy of Russification during the Soviet period, the revival of Kazakh language and culture, and the influence of global digital media all shape the current educational en-

vironment. In this context, identity formation becomes both a challenge and an opportunity. Government programs such as *Rukhani Zhangyru* (Spiritual Modernization) (Nazarbayev, 2017) emphasize the need to combine national heritage with preparation for global citizenship.

One useful way to achieve this balance is through corpus-based resources – digital collections of texts that can be studied with linguistic software. Such resources help both teachers and students to see how language carries cultural meaning, ethical values, and historical perspectives. For Kazakhstan, the *Abai Poetic Corpus* (*Abai Poetic Corpus*) is especially valuable. It provides a fully annotated, searchable collection of Abai Kunanbayev's poetry. With tools such as lemma search, keyword-in-context displays, frequency lists, and collocation analysis, learners can explore Abai's language in systematic and interpretive ways.

The use of corpora in language education has grown worldwide in recent decades. As Hunston (2002) observed, corpora influence teaching in two main ways: they give us new, evidence-based descriptions of language and they act as direct resources for developing materials and shaping methods. Both functions are now relevant for digital linguo-

didactics and for integrating literary heritage into education.

Abai Kunanbayev (1845-1904) is regarded as a central figure of Kazakh cultural identity. His writings combine philosophy and poetry, offering an ethical vision of personal growth and civic duty. A key idea in his work is *adam bolu*, becoming a true human being – which is achieved through the balance of *aqyl* (intellect), *qairat* (willpower), and *zhurek* (heart, compassion). His worldview draws on Islamic humanism, moral reflection, and Enlightenment thought, creating a holistic model of ethical citizenship that still speaks to issues of youth identity, social unity, and moral education today.

As Makatova (2020) notes, Abai's works guide the spiritual growth of the nation and provide tools for addressing ethical and cultural problems. His poetry is more than art; it acts as a moral reference point for social change. Yet in schools, Abai is often taught in a traditional way, emphasizing memorization instead of interpretation. Many students see his words as disconnected quotations rather than as part of a broader moral message. A corpus-based approach, by contrast, invites learners to examine Abai's language themselves, notice patterns, and build their own understanding of his values. This approach fits well with modern educational goals such as critical thinking, digital literacy, and identity development.

This study investigates the Abai Poetic Corpus as a digital tool for both language education and national identity formation. It focuses on how corpus-based methods can help students better understand the moral concepts in Abai's poetry and engage more thoughtfully with national values. By combining linguistic analysis with classroom practice in schools and universities, the research adds to current work on corpus-informed education. The ultimate aim is to show how digital resources can bring classical literature to life and make cultural heritage more relevant, engaging, and transformative for younger generations.

Materials and methods

This study used a mixed-methods design, combining digital text analysis with a classroom experiment to see how Abai's poetry can help students connect with national identity and moral values in today's digital environment. The project had two stages: (1) a linguistic study of the Abai Poetic Corpus with digital tools, and (2) an educational trial with secondary school and university students.

1. Corpus and Tools

The main resource was the *Abai Poetic Corpus* (*Abai Poetic Corpus*), an open-access platform that contains Abai Kunanbayev's poetry and prose in Kazakh, annotated for language and meaning. For the linguistic part, AntConc (v.3.5.9) was applied to identify culturally important words through frequency counts, explore collocations of moral terms, and check real examples in concordance lines. The key concepts selected for analysis were *adam* (human), *aqyl* (intellect), *iman* (faith/conscience), *namys* (honor), *qanagat* (contentment), and *qazaq* (Kazakh).

2. Educational Pilot

To test how the corpus could be used in teaching, a pilot study was carried out with 60 participants: 30 Grade 10 pupils (aged 15-16) from a Kazakh-medium school and 30 first-year university students in the humanities.

The intervention ran across three one-hour sessions:

1. A pre-task survey measured students' baseline knowledge of national and ethical values;
2. Learners then worked with a simplified version of the Abai Corpus to search for words, explore collocations, and discuss meanings in context;
3. Finally, a post-task survey and open-ended reflections captured their impressions and insights.

Secondary students mainly worked in groups, while university students explored the corpus individually, with all tasks conducted in Kazakh. Teachers supported the activities with both written and oral instructions.

3. Data Collection and Analysis

Data included both numbers and narratives.

Quantitative data: pre- and post-survey responses (Likert scales, multiple-choice), and accuracy in recognizing Abai's value triad (*aqyl–qairat–zhurek*).

Qualitative data: students' written reflections, classroom observation notes, and concordance lines chosen for discussion.

Survey results were analyzed with descriptive statistics to track improvements, while thematic coding was used for reflections to reveal how learners interpreted values and engaged emotionally with Abai's ideas.

Literature review

In recent years, researchers have shown increasing interest in using digital corpora for both educational and linguistic purposes, particularly when it

comes to questions of national identity. In Kazakhstan, the *Abai Poetic Corpus* hosted on the national platform Corpora.kz has become an important tool for studying Kazakh literary heritage through modern, data-driven methods (Baimyrza et al., 2024a). The corpus not only makes Abai Kunanbayev's poetry available in digital and annotated form, but also provides tools such as morphological tagging, concordance searches, and keyword indexing. This makes it a valuable resource for digital linguodidactics.

Methodologically, the design of the Abai Poetic Corpus follows international practices in corpus linguistics. It involved careful curation of data, detailed metadata annotation, and the creation of user-friendly search functions. As described in Baimyrza et al. (2024b), the team faced both technical and conceptual challenges in the process, but their goal was to create a resource not only for researchers, but also for teachers and students. This reflects the broader "pedagogical turn" in corpus linguistics, noted by Flowerdew (2009), where authentic language data is brought directly into classrooms to reshape both teaching content and methodology.

Corpus-based analysis has also been applied to the poetic and rhetorical features of Abai's work. For example, Adilov and Petek (2022) show how tagging and automatic annotation allow large-scale analysis of verse structures – meter, rhyme, and rhetorical devices, which previously depended on close reading alone. This makes it possible to study rhythm, symmetry, and stylistic patterns in a more systematic and replicable way.

Work has also been done on historical phonology and orthographic variation. Comparing early manuscripts with modern editions of Abai's poetry reveals changes in spelling, phonetic representation, and dialectal features (Adilov & Petek, 2022). Such studies contribute to debates about textual authenticity and the development of standard Kazakh.

At the same time, Abai's poetry invites meta-poetic interpretation, where the poet reflects on his own craft. Aimukhambet (2021) situates Abai in this framework, noting how his voice is both reflective and instructional. Through corpus analysis, elements such as metaphor density, poetic syntax, and reflexive phrasing can be studied systematically, providing insights into how Abai constructed his identity as a poet and thinker.

Lexico-semantic and metric studies also reveal how Abai used synonymy, antonymy, and polysemy to convey ethical and philosophical meaning. Pirmanova et al. (2024) demonstrate how collocation

patterns and rhyme schemes highlight the coherence and richness of Abai's poetic thought.

Beyond literary research, corpora are increasingly used in language pedagogy. Aston (2000) emphasizes that corpora promote learner autonomy by encouraging students to discover patterns themselves rather than rely only on rules. Mukherjee (2006) adds that corpora connect theory and practice, serving both as a source of linguistic description and as a teaching resource. For Kazakhstan, these approaches are particularly relevant for modernizing philological education.

One of the most promising uses of the Abai Poetic Corpus is in philology programs. Pirmanova et al. (2025) argue that discovery-based learning (DDL) using the corpus fosters critical thinking, stylistic awareness, and digital literacy. By testing hypotheses and analyzing patterns, students become active participants in interpreting Abai's texts. This fits well with the calls by Aston (2000) and Flowerdew (2009) for corpus-informed, student-centered curricula.

Of course, challenges remain. Madiyeva et al. (2025) points out barriers such as limited teacher training, lack of digital resources, and the persistence of traditional teaching methods. They suggest teacher workshops, modular digital tools, and the integration of corpus literacy into teacher training programs as possible solutions.

The Abai corpus also has broader applications for literary history, including questions of authorship, manuscript provenance, and textual transmission. Corpus-based comparisons can help clarify which poems genuinely belong to Abai, contributing to ongoing debates in Kazakh literary scholarship.

Taken together, these studies show the broad potential of the Abai Poetic Corpus as both a literary and educational tool. They underline how corpus-based approaches, when applied to culturally significant texts, can strengthen identity-focused education, support critical engagement, and modernize philological training in Kazakhstan. At the same time, there is still a lack of empirical classroom-based studies. This research seeks to address that gap by exploring how the Abai Poetic Corpus can be applied in digital linguodidactics to support culturally grounded and ethically meaningful education.

Results and discussion

The study's findings are organized into three parts: (1) linguistic insights from corpus analysis; (2) student learning outcomes following the inter-

vention; and (3) a comparison between secondary and university participants.

1. Corpus Patterns of National and Moral Values

Exploration of the *Abai Poetic Corpus* through AntConc revealed a network of culturally loaded and morally significant words central to Kazakh thought. The most salient terms included: *adam* (human) – 152 instances, *aqyl* (intellect/wisdom)

– 88, *iman* (faith/conscience) – 41, *namys* (honor/dignity) – 46, *qanagat* (contentment/moderation) – 27, and *qazaq* (Kazakh) – 62. Collocational patterns showed that these notions often co-occurred with verbs like *bilu* (to know), *tusinu* (to understand), and *bagalau* (to value), underscoring their instructional and ethical role within Abai's language.

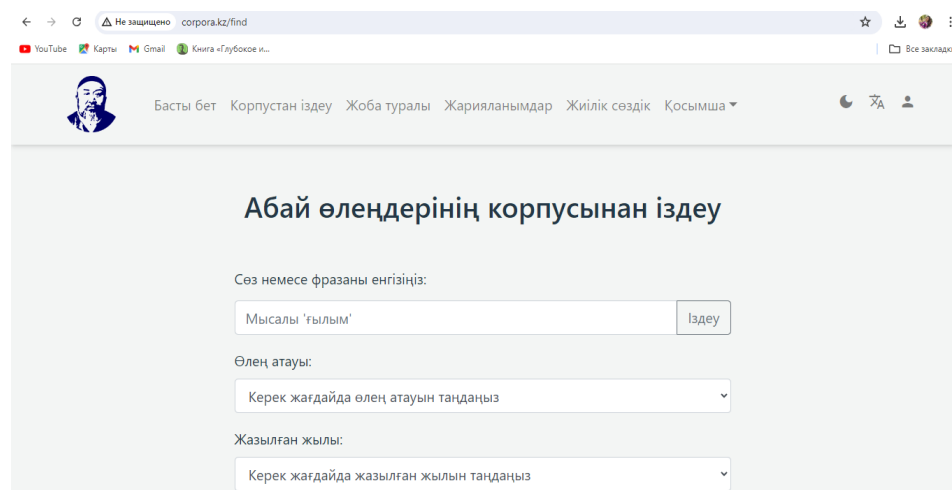


Figure 1 – A searching page of Abai Poetic Corpus at www.corpora.kz

The triad of *aqyl*, *qairat*, and *zhurek* emerged consistently in both verse and prose, highlighting its status as a moral and philosophical foundation. Abai's idea of the *tolyq adam* (complete person) – one who harmonizes intellect, will, and heart – reflects a synthesis of Kazakh ethical ideals and Islamic humanist traditions.

2. Student Learning Outcomes

The pilot intervention produced measurable improvements in both groups.

1. “Before, I only memorized Abai’s poems. Now I feel he speaks directly about how to live well”.

2. “I never realized ‘iman’ implies responsibility to oneself and others. It feels like a guide for being Kazakh”.

3. “Abai’s *aqyl*, *qairat*, *zhurek* helped me see the balance we need in life”.

Secondary students tended to express ideas in emotional and concrete terms, while university learners showed a greater inclination to connect values to abstract philosophical and civic frameworks.

Table 1 – Quantitative Results

Indicator	Pre-task (%)	Post-task (%)
Correct definition of <i>namys</i> , <i>iman</i> , <i>aqyl</i>	38%	73%
Recognition of Abai’s moral triad	27%	70%
Increased interest in Abai’s work (self-reported)	–	84%

The data demonstrate a significant gain in learners’ conceptual understanding after working with corpus tasks. Qualitative reflections added further depth:

Table 2 – Comparative Reflections

Category	Secondary Students	University Students
Engagement	High with visual/concrete tasks	Moderate, but independent
Interpretation	Emotionally expressive	Abstract and analytical
Vocabulary use	Emerging, paraphrased	Precise, contextually accurate
Collaboration	Group-oriented	Reflective, often individual

These differences align with Vygotsky's (1978) view of scaffolding across developmental stages, confirming that corpus-based tasks can be adapted to varied learner needs by adjusting complexity.

The intervention showed that:

- Students achieved a deeper grasp of moral and national values embedded in Abai's texts;
- Corpus-informed activities fostered both emotional engagement and critical reflection;
- The digital format enhanced accessibility and learner motivation;
- The approach can be successfully scaled across age levels and institutions.

Taken together, the findings validate the claim that integrating corpus linguistics with culturally resonant materials such as Abai's works strengthens national identity education in Kazakhstan.

The findings of this study point to the powerful role that corpus-based methods can play in connecting language learning with ethical reflection and cultural identity. By working with the *Abai Poetic Corpus*, students were not only practicing digital literacy but also engaging with value-based education – a combination that is highly relevant for Kazakhstan's multilingual and post-Soviet context.

1. Abai's Moral Language as a National Resource

Analysis of Abai's poetry shows that his moral and philosophical system is both consistent and teachable. Words like *adam* (human), *aqyl* (intellect), *iman* (faith/conscience), *namys* (honor), and *qanagat* (contentment) appear frequently and in meaningful contexts. This supports Zhubanov's (2019) claim that Abai's vocabulary functions as a code of indigenous ethical thought. Just as importantly, these terms often appear with verbs such as *bil* (know), *tusin* (understand), *bagala* (value), *iste* (do), showing that Abai's ethics are active and practical, not abstract. His model of the *tolyq adam* (complete human) – balancing intellect, will, and compassion – remains both a psychological and pedagogical framework for identity formation. Through corpus analysis, students can see for themselves how language encodes moral ideas, which encourages not just awareness of words but also deeper reflection and a stronger sense of belonging.

2. Learner Engagement Through Corpus Tools

The teaching experiment showed clear benefits when students worked directly with the corpus. Test results indicated better recognition and use of key moral terms, while interviews revealed that students started to see Abai's texts as more than exam material. They described them as living sources of guid-

ance for understanding what it means to be Kazakh. These outcomes echo findings by Braun (2005), Boulton (2010), and Gablasova et al. (2017), who argue that data-driven learning (DDL) supports autonomy, critical thinking, and deeper comprehension. Interestingly, younger students responded best to visual concordance displays and teacher guidance, while university students preferred more independent analysis. This pattern reflects Vygotsky's (1978) idea of the Zone of Proximal Development, showing that tools must be adapted to learners' stages and capacities.

3. Reimagining the Teaching of Abai

In many classrooms, Abai is still taught in a traditional way, with emphasis on biography and history rather than his ethical ideas (Ilyasova & Tazhibayeva, 2020). This often leads to rote memorization instead of personal engagement. The corpus-based approach changes this dynamic: it turns students into investigators, prompting discovery and dialogue. This reflects Halstead and Taylor's (2000) call for moral education that fosters personal meaning-making, and Banks' (2008) argument that identity education should be authentic but also open to student voice. For Kazakhstan, where modernization and tradition must coexist, this shift is more than pedagogy – it is a cultural necessity.

4. National Identity in a Digital Age

Digitizing Abai's works creates new opportunities for young people to engage with heritage. Today's students navigate competing identity narratives – Kazakh, Russian, Turkic, Islamic, Western, and need tools that ground them without imposing ideology. The Abai Digital Corpus offers such a platform: it revitalizes a cultural legacy in digital form, encouraging students to construct identity through language exploration and reflection. This aligns with the aims of *Rukhani Zhangyru*, which seeks to balance innovation with tradition. More broadly, the study shows that corpus linguistics, when applied to culturally meaningful content, can support not only language education but also identity building, ethical inquiry, and civic growth. In this way, corpus tools act both as mirror of the past and as frameworks for shaping the future.

Conclusion

This study set out to explore how the Abai Poetic Corpus can function as a digital tool for teaching national identity and ethical literacy among Kazakhstani students. Positioned at the crossroads of digital linguodidactics, corpus-based learning, and values education, the project showed that digital corpora

are able to give classical literature new relevance in today's classrooms.

The results indicate that Abai's poetic language, when examined through interactive corpus tools, provides learners with a more direct and personal connection to national values. Key concepts such as *adam bolu* (being human), *aqyl* (intellect), *iman* (faith), *namys* (honor), and *qanagat* (contentment) were not only identified but also reinterpreted by students in light of their own social experiences. The analysis showed that both school and university students became more aware of the ethical triad of *aqyl – qairat – zhurek*, and that corpus-supported activities could engage both their reasoning and their emotions.

Comparisons between the two groups revealed differences in interpretation, vocabulary, and styles of reflection. Nevertheless, across the board, corpus tasks encouraged curiosity, critical thinking, and a culturally grounded type of literacy. Students themselves noted that the digital format made Abai's poetry feel «alive», «relevant», and «personally significant», in contrast to the more passive memorization usually linked with classical texts.

From a teaching perspective, the research confirms that integrating national literary heritage into digital environments is not only possible but also beneficial. It supports broader goals of developing independent learners who are confident in their cultural identity and at the same time comfortable with digital technologies. In this way, the Abai Poetic Corpus illustrates how digitized literature can be used for reflective, interdisciplinary, and identity-focused education.

The broader implications are also notable. This work adds to discussions on culturally responsive pedagogy in postcolonial and multilingual settings. For Kazakhstan, it provides a concrete example of how digital humanities can advance state priorities such as *Rukhani Zhangyru* by linking tradition with modern innovation. It also points to future directions in curriculum design, teacher preparation, and the creation of digital corpora for other Kazakh writers and genres.

Future studies might build on these findings by tracing the long-term effects of corpus-based teaching, testing combinations with AI-powered language tools, or examining how Abai is received by students from different linguistic communities. Teacher perspectives and alignment with national curricula also require further attention to make the integration sustainable.

In conclusion, the Abai Poetic Corpus should not be seen simply as a digital archive. It works as an active educational tool that helps young people connect with their cultural roots, develop ethical reasoning, and find their own voice in a rapidly changing society. Its use in digital linguodidactics marks a step toward a more reflective, engaged, and culturally grounded model of education in Kazakhstan.

This research is funded by the Committee of Science of the Ministry of Science and Higher Education of the Republic of Kazakhstan (Grant No.BR21882334 «Kazakh poetic corpus development: morphological and poetic designation of Abai's poems»).

References

- Abai Poetic Corpus. [Electronic Resource]. URL: <http://corpora.kz/> (Date of use: 21.12.2024)
- Adilov M., & Petek E. (2022). Language of Abai poems manuscripts (some phonetic differences). *Turkic Studies Journal*, 4(1), 70–85. <https://doi.org/10.32523/2664-5157-2022-1-70-85>
- Aimukhambet Zh. (2021). Abai olenderi metapoetica paradigmasynnda (Abai's poems in the paradigm of metapoetics). *Bulletin of L.N. Gumilyov Eurasian National University. Philology Series*, 136(3), 84–93. <https://doi.org/10.32523/2616-678X-2021-136-3-84-93> (In Kazakh)
- Aston G. (2000). Corpora and language teaching. In L. Burnard & T. McEnery (Eds.), *Rethinking language pedagogy from a corpus perspective* (pp. 7–17). Frankfurt: Peter Lang.
- Baimyrza A., Pirmanova K., & Serikbayeva A. (2024a). Challenges and prospects of creating Abai poetic corpus. *Bulletin of the Karaganda University. Philology Series*, 56–66. <https://doi.org/10.31489/2024Ph4/56-66>
- Baimyrza A., Serikbayeva A., Pirmanova K., Pirmanova A., & Myrzakhanova G. (2024b). The impact of linguistic identity on the perception of Abai's poetry. *Forum for Linguistic Studies*, 6(5), 239–246. <https://doi.org/10.30564/fls.v6i5.6879>
- Flowerdew L. (2009). Applying corpus linguistics to pedagogy: A critical evaluation. *International Journal of Corpus Linguistics*, 14(3), 393–417. <https://doi.org/10.1075/ijcl.14.3.05flo>
- Hunston S. (2002). *Corpora in applied linguistics*. Cambridge: Cambridge University Press.
- Madiyeva G., Akhmed-Zaki D., Mansurova M., Kyrgyzbayeva M., & Kadyrbek N. (2021). Development of the information system for the Kazakh language preprocessing. *Cogent Engineering*, 8(1), 1896418. <https://doi.org/10.1080/23311916.2021.1896418>

Madiyeva G. B. (2024). Innovations in studying Kazakh poetry: From theory to practice in developing the Kazakh poetic corpus. In Proceedings of the International Scientific-Practical Conference “Modern Technologies in Computational Linguistics – CTCL.2024” (April 26, 2024). Tashkent, Uzbekistan.

Makatova A. (2020). Abai murasy – ruhani zhangyru ozegi (Abai’s legacy is at the core of spiritual modernization). [Electronic Resource]. URL: <https://abaialemi.kz/kz/post/view?id=1887> (Date of use: 21.12.2024) (In Kazakh)

Mukherjee J. (2006). Corpus linguistics and language pedagogy: The state of the art – and beyond. English Corpus Linguistics, Vol. 3, P. 5-24.

Nazarbayev N.A. (2017). President’s article “Course towards the Future: Modernization of Kazakhstan’s Identity”. [Electronic Resource]. URL: https://www.akorda.kz/kz/events/akorda_news/press_conferences/memleket-basshysynyn-bolashakka-bagdar-ruhani-zhangyru-atty-makalasy (Date of use: 21.12.2024)

Tatieva S. S., Kazezova A. M., & Ongalbek A. A. (2024). An integrative approach to developing the Kazakh small poetic corpus of Abai: Prospects for research in Kazakh philology. In Proceedings of the VIII International Scientific and Practical Conference “Science in Modern Society” (April 30 – May 1, 2024). Beijing, China.

Tokmyrzaev D. O., Zhubanova E. A., & Pirmanova A. K. (2024). Creating a probabilistic-statistical model of the Kazakh text. In Proceedings of the International Scientific-Practical Conference “Modern Technologies in Computational Linguistics – CTCL.2024” (April 26, 2024). Tashkent, Uzbekistan.

Zaurebekova G. O., Ayazbekova K. A., & Ongalbek A. A. (2024). Initiatives for developing the Kazakh poetic corpus based on Abai’s poems. In Proceedings of the International Conference “Theoretical Hypotheses and Empirical Results” (March 28–29, 2024). Oslo, Norway.

Information about author:

Baimyrza Ainur – PhD Student, A. Baitursynuly Institute of Linguistics (Kazakhstan, Almaty, e-mail: baimurzina.ainur@gmail.com).

Автор туралы мәлімет:

Баймырза Айнұр – PhD докторант, А. Байтұрсынұлы атындағы Тіл білімі институты (Қазақстан, Алматы қ., e-mail: baimurzina.ainur@gmail.com).

Сведения об авторе:

Баймырза Айнұр – PhD докторант, Институт языкознания им. А. Байтұрсынұлы (Алматы, Казахстан, e-mail: baimurzina.ainur@gmail.com).

Date of receipt of the article: June 07, 2025.

Accepted: September 22, 2025.