ARTICLE

Educational Potential of Abai's Corpus in Developing National Identity in the Digital Age

Gulmira Madiyeva [iD] *, Ainur Baimyrza* [iD] *, Kunsulu Pirmanova** [iD] *, Aizhan Serikbayeva* [iD] *,Darkhan Tokmyrzayev* [iD]

Department of Turkology and Language Theory, Al-Farabi Kazakh National University, Almaty 050000, Kazakhstan

ABSTRACT

This study examines the potential of the *Abai Digital Corpus* to foster national identity and moral development among secondary and university students in Kazakhstan. Drawing on the culturally and ethically rich texts of Abai Kunanbayev, the research integrates corpus linguistics methods with a classroom-based pedagogical intervention. The study identifies and analyzes key concepts such as *namys* (honor), *aqyl* (intellect), *zhurek* (heart), and *iman* (faith), which are foundational to Abai's philosophy of the *tolyq adam* (complete person). These values were embedded into interactive, task-based learning modules designed to stimulate ethical reflection, cultural engagement, and critical thinking. The methodology included triangulated data collection through pre-and post-task surveys, participant reflections, and classroom observation. Using AntConc software, the corpus analysis revealed meaningful patterns in students' engagement with moral terms and their associations. The findings suggest that a corpus-informed approach not only improves students' digital and linguistic skills but also deepens their connection to national heritage and identity. Moreover, this study underscores the relevance of Abai's thought in contemporary education and offers a scalable model for integrating digital humanities tools into value-based education. It aligns with Kazakhstan's *Rukhani Zhangyru* initiative, emphasizing cultural modernization through education. Recommendations for future research include extending the age range of participants and incorporating longitudinal analysis to assess the sustained impact of such interventions.

Keywords: Poetic Corpus; Abai Kunanbayev; National Identity; Digital Age; Education; Abai's Poetry

1. Introduction

In the rapidly evolving landscape of 21st-century education, there is a growing imperative to transcend traditional models focused solely on the transmission of factual knowledge. Modern pedagogy increasingly emphasizes the cultivation of values-based learning, critical thinking, emotional intelligence, and the formation of personal and collective identity. In this context, education is seen not only as a means of preparing individuals for the labor market but also as a key instrument in shaping ethical citizens and cohesive societies. For countries undergoing complex socio-cultural transitions—such as Kazakhstan, a multilingual and multiethnic post-Soviet state—the role of education in fostering a coherent and confident national identity among young people is both urgent and multifaceted.

Kazakhstan's post-independence era has been marked by significant efforts to reclaim and reconstruct cultural heritage, assert linguistic sovereignty, and reestablish national narratives rooted in indigenous knowledge systems. Amid these processes, the education system plays a crucial role in transmitting historically grounded yet future-oriented cultural values. Central to this mission is the integration of local intellectual traditions and ethical frameworks into the curriculum. However, challenges persist: competing global influences, linguistic hybridity, and the remnants of Soviet-era ideologies often dilute the presence and impact of authentic Kazakh voices in mainstream education.

Among the most enduring and influential sources of Kazakh national consciousness is the legacy of Abai Kunanbayev (1845–1904), whose poetic and philosophical writings offer a profound synthesis of moral introspection, cultural wisdom, and intellectual inquiry. Abai's work articulates essential themes such as justice, human dignity, education, personal responsibility, and the pursuit of truth—values that continue to resonate with the ethical and civic aspirations of contemporary Kazakh society. His unique position as both a critic of societal shortcomings and a visionary thinker bridging Eastern spiritual traditions with Western rationalism renders his texts exceptionally relevant for education that seeks to nurture the *tolyq adam* (complete person)—a morally and intellectually developed individual.

With the emergence of the Abai Digital Corpus, a new dimension of pedagogical engagement has become possible. This structured, annotated, and searchable collection of Abai's poetic and prose works offers students and educators an innovative tool for exploring language, culture, and identity through a data-driven lens. Corpus-based learning methods enable students to interact directly with authentic texts, uncovering recurring patterns, thematic clusters, and key moral concepts embedded in the language. This process not only enhances digital literacy and critical text analysis skills but also fosters a deeper emotional and intellectual connection to national heritage.

This paper proposes that the integration of the Abai Digital Corpus into formal education—particularly within humanities and elective subjects such as Abaitanu (Abai Studies)—can play a transformative role in shaping students' understanding of national identity. Through thoughtfully designed, corpus-informed tasks—including semantic analysis, conceptual mapping, reflective writing, and dialogic interpretation—students can engage with Abai's texts in ways that are personally meaningful and culturally empowering. Moreover, this approach supports Kazakhstan's broader strategic goals, including the "Rukhani Zhangyru" (Spiritual Modernization) initiative, the digital transformation of education, and efforts to decolonize curricula by re-centering indigenous epistemologies.

The central aim of this study is to investigate the educational potential of the Abai Digital Corpus in promoting national identity among secondary and university-level students in Kazakhstan. Specifically, it explores how corpus-informed educational interventions can support the development of cultural consciousness, moral reasoning, and emotional engagement with Kazakh heritage in a generation that is increasingly digitally fluent and globally connected.

Research Questions

1. How can the Abai Digital Corpus be effectively integrated into secondary and higher education curricula to support the development of national identity?

2. What specific national values and cultural concepts can be identified in Abai's works through corpus-based analysis?

3. To what extent do corpus-informed learning tasks enhance students' awareness, reflection, and inter-

pretation of Kazakh national identity?

4. How does the use of digital tools and corpus linguistics contribute to the modernization and relevance of Abai studies in contemporary education?

2. Literature Review

National identity is a multifaceted and evolving construct that encompasses shared language, cultural practices, historical memory, moral frameworks, and a sense of belonging to a particular community [1]. In the realm of education, national identity is not only a curriculum goal but also a dynamic process of socialization through which learners internalize the values, narratives, and linguistic norms of their community. This process becomes especially complex in multicultural, postcolonial states such as Kazakhstan, where the intersections of ethnicity, language policy, and geopolitical influence continue to shape public discourse and educational policy [2].

Since gaining independence in 1991, Kazakhstan has undertaken significant reforms to reclaim its national heritage and rebuild a collective identity rooted in indigenous epistemologies. The *Rukhani Zhangyru* (Spiritual Modernization) program launched in 2017 represents a key policy initiative aimed at strengthening cultural consciousness, promoting historical awareness, and preparing youth to be modern, yet culturally grounded, citizens. This program emphasizes the importance of returning to foundational national values, reinvigorating traditional knowledge systems, and integrating them with digital and global competencies.

According to Banks [3], effective national identity education must be inclusive, dialogical, and grounded in the authentic cultural experiences of learners. He argues that students need to see their histories, languages, and belief systems reflected meaningfully in the curriculum in order to internalize a sense of belonging and shared purpose. Halstead and Taylor [4] similarly assert that moral and values education must be sensitive to local cultural contexts, balancing universal ethical principles with traditions specific to a community's worldview. These arguments are particularly relevant in Kazakhstan, where youth often negotiate competing identities— Kazakh, Russian, and global—in their everyday lives.

The tension between linguistic diversity and national unity is one of the most pressing educational challenges in Kazakhstan. Suleimenova et al. [5] emphasize that language remains central to identity development, and multilingualism must be carefully managed to avoid eroding native linguistic and cultural resources. Despite the official promotion of Kazakh as the state language, Russian remains dominant in many spheres of education and media. As such, there is a growing need for pedagogical strategies that make Kazakh-language content not only accessible but also engaging and personally meaningful for students.

One promising path toward addressing this challenge lies in the works of Abai Kunanbayev, whose poetry and philosophical prose have long been considered the moral and intellectual foundation of Kazakh national identity. His *Qara Sozder* (Words of Edification) and lyrical poems advocate for personal integrity, lifelong education, ethical reasoning, and spiritual development. Abai's vision of the *tolyq adam*—the complete human being— encompasses the harmonious integration of *aqyl* (intellect), *qairat* (willpower), and *zhurek* (heart), offering a holistic ethical framework that continues to resonate with contemporary Kazakh values [6].

The enduring educational relevance of Abai has been extensively documented. M.O. Auezov [7], in his seminal biography *Abai zholy* (The Path of Abai), presents Abai not only as a cultural icon but also as an educational reformer whose vision sought to uplift the Kazakh people through reason, literacy, and self-awareness. Abai's writings promote civic responsibility, moral introspection, and social critique—traits essential for fostering a thoughtful and engaged citizenry.

Linguist A. Zhubanov [8] extends this interpretation by analyzing how Abai's lexicon encodes national values at the structural level of language, making his works ideal for teaching identity through linguistically rich content. However, as Ilyasova and Tazhibayeva [9] observe, Abai's pedagogical potential remains underutilized in current practice, which often relies on rote memorization and textbook-driven instruction. Students may learn to recite Abai's poems, but they are rarely invited to explore the deeper moral and philosophical layers of his writing. To make Abai's legacy more impactful, new pedagogical approaches are needed— ones that prioritize interpretation, personal reflection, and

digital interactivity.

This is where corpus linguistics offers transformative potential. Corpus-based methods, which involve analyzing patterns of language across large text collections, allow students to engage with literature as both readers and researchers. By identifying key concepts, collocations, and thematic clusters, learners can uncover how values are constructed and conveyed through language [10]. These methods are particularly effective when applied to culturally significant texts, as they help bridge the gap between language form and cultural meaning.

Baker [11] and Braun [12] have shown that corpora enhance critical discourse analysis and promote learner autonomy by inviting students to make their own observations and interpretations. Boulton [13] argues that data-driven learning (DDL) fosters critical thinking and discovery-based engagement, particularly in tasks involving reflective writing and close reading. When applied to literature education, corpus methods enable students to uncover hidden patterns and connect with texts on a conceptual level.

Gablasova, Brezina, and McEnery [14] emphasize that corpus tools do more than improve linguistic competence—they can also support identity formation. When students examine culturally resonant texts like those of Abai, they gain insight into how moral values, worldviews, and historical narratives are encoded in language. This creates a powerful space for personal and cultural reflection, especially for youth navigating issues of belonging and cultural continuity.

In recent years, the educational potential of Abai's poetic corpus has been increasingly explored through interdisciplinary lenses, particularly at the intersection of digital humanities, corpus linguistics, and identity formation. Building on prior foundational research, recent studies have significantly enriched the field. For example, Baimyrza et al. [15] examined how linguistic identity shapes students' perception of Abai's poetry, revealing that language background plays a key role in emotional and cognitive engagement with the texts. Similarly, Hanauer [16] provided both theoretical and empirical insights into how poetry functions within the learning process, emphasizing its affective and reflective dimensions. Extending this perspective, Hani et al. [17] investigated the role of poetry in linguistic education, affirming its value as a culturally embedded and cognitively stimulating pedagogical tool.

The integration of Abai's poetry into digital formats has also given rise to technical and philological innovations. Madiyeva [18] proposed practical strategies for developing the Kazakh poetic corpus, while Ayazbekova, Zaurbekova, Tatieva, and Kazezova [19] offered an integrative approach grounded in Kazakh philological traditions. Technical implementations have been supported by work on preprocessing and statistical modeling of Kazakh texts [20,21], enhancing the computational accessibility of Abai's works. Additionally, Zaurbekova et al. [22] highlighted grassroots initiatives aimed at corpus development based on Abai's poems, reflecting growing institutional interest in his digital legacy.

These studies are complemented by more philosophical and sociolinguistic perspectives. Severskaya [23] emphasized the contextual nature of poetic meaning-making in sociolinguistic poetics, while Yessyim [24] approached Abai's work through the lens of his moral and intellectual teachings—central to the formation of national identity. Together, these contributions underscore the multidimensional value of Abai's corpus not only as a literary archive but as a culturally formative tool in contemporary education.

While corpus linguistics is still an emerging field in Kazakhstan's educational system, the development of the Abai Digital Corpus marks a significant step forward. This tool provides teachers and students with an accessible, annotated, and searchable collection of Abai's works. It enables semantic analysis of core national concepts such as *adam* (human), *iman* (faith), *namys* (honor), and *qanagat* (contentment), and offers a foundation for designing interactive and meaningful learning activities. By linking linguistic patterns to ethical themes, the corpus allows students to experience Abai's philosophy as a living discourse rather than a static legacy.

In sum, the reviewed literature affirms that national identity education is most effective when it is culturally rooted, interpretive, and participatory. Abai's works, rich in linguistic, ethical, and philosophical content, provide an ideal foundation for this kind of education. When paired with corpus linguistics, they offer a unique opportunity to merge tradition with innovation, enabling students

to explore national identity not as something abstract, but as something they actively construct and internalize through language. The integration of the Abai Digital Corpus into educational practice represents a timely and evidence-based strategy to modernize Kazakhstani pedagogy while preserving and revitalizing its cultural heritage.

3. Methodology

This study adopts a mixed-methods approach combining corpus-based linguistic analysis with an educational intervention aimed at fostering national identity among Kazakhstani students. The methodology is structured into four components: (1) corpus analysis, (2) participants, (3) instructional procedure, and (4) data collection and analysis.

3.1. Corpus-Based Linguistic Analysis

The primary data source was the Abai Digital Corpus, which contains annotated versions of Abai Kunanbayev's poetic and prose works. Using the AntConc software, a series of analyses were conducted to identify high-frequency moral and national identity-related terms, such as adam (human), aqyl *(*intellect), iman (faith), and namys (honor). Collocational and concordance patterns were examined to understand how these concepts are linguistically constructed and morally framed within Abai's discourse.

The results of this analysis (see Results section) informed the design of targeted pedagogical tasks that integrated corpus data into classroom activities.

3.2. Participants

The study involved two groups of participants: 1) 30 secondary school students from Grade 10 (ages 15–16), attending a public school in urban Kazakhstan; 2) 30 university students in their first year of a humanities program.

All participants were native or fluent Kazakh speakers with prior exposure to Abai's works through the national literature curriculum. Participation was voluntary, and all responses were anonymized for ethical compliance.

3.3. Instructional Procedure

The intervention consisted of a single 90-minute classroom session for each group, designed to implement corpus-informed learning tasks (see Appendix A).

The procedure included:

a) Orientation (10 minutes)

Brief introduction to Abai's moral philosophy and key concepts. Demonstration of corpus tools and concordance reading.

b) Corpus-Based Task Session (60 minutes)

Students completed four structured learning tasks focusing on: concordance analysis (*iman*, *namys*); concept mapping (*tolyq adam* model); interpreting keyword frequency; identity-based reflective writing. Tasks were completed individually, with collaborative discussion encouraged during the final stages.

c) Reflection and Feedback (20 minutes)

Students answered open-ended questions and participated in a class discussion. Facilitators collected written reflections and observation notes.

The session design was informed by previous research in corpus-based pedagogy [12,13] and culturally responsive learning [3].

3.4. Data Collection and Analysis

Three types of data were collected: qualitative, quantitative, observational data**.** Written reflections, concept maps, and student answers to corpus tasks were analyzed thematically, focusing on how students engaged with moral values and linguistic expressions of national identity (using Braun & Clarke's method of thematic analysis [12]). Pre- and post-task surveys measured change in understanding of key concepts. Frequency counts and descriptive statistics were used to compare results. Field notes recorded during task execution and discussion captured patterns of interaction, engagement, and critical thinking.

This multi-source triangulation enhanced the reliability of the findings and allowed for a nuanced understanding of how corpus-based education can influence identity formation.

4. Results

The results of this study are presented in three subsections: (1) findings from the corpus-based linguistic analysis of Abai's works, (2) outcomes of the educational

intervention conducted among secondary and university students, and (3) comparative insights across the two learner groups.

4.1. Corpus-Based Identification of National Values

The corpus analysis of Abai's poetry and *Qara Sozder* revealed a recurring network of culturally and morally significant lexicon. Utilizing AntConc software, keyword frequency, collocation, and concordance line analysis were applied to detect key concepts aligned with Kazakh national identity.

The word *adam* (human) occurred 152 times, often framed in moral contexts emphasizing dignity, wisdom, and self-improvement.

Other central values included: *aqyl* (intellect)—88 occurrences, *iman* (faith/conscience)—41 occurrences, *namys* (honor)—46 occurrences, *qanagat* (contentment)—27 occurrences, *qazaq* (Kazakh)—62 occurrences.

Concordance line analysis showed these terms often appeared in proximity to action-oriented verbs such as *bil* (know), *tusin* (understand), and *bagala* (value), reinforcing their role as actionable moral guidance [6].

The triadic model *aqyl - qairat - zhurek*, referenced in both poetic and prose texts, emerged as a cohesive ethical structure for personal and social conduct. This confirms Zhubanov's [8] observation that Abai's language encodes an indigenous moral philosophy suitable for value-based pedagogy.

4.2. Educational Intervention Results

To assess the pedagogical potential of the Abai Digital Corpus, a pilot study was conducted involving:

- 30 Grade 10 students (age 15–16) in a secondary school;
- 30 first-year humanities students in a public university.

Students engaged in three sessions involving: pre-task survey, corpus-based reading and reflective tasks, post-task survey and open-ended writing (**Table 1**).

Table 1. Quantitative findings.

Indicator	Pre-task (%)	Post-task (%)
Correctly defined *namys*, *iman*, *aqyl*	38%	73%
Recognized Abai's moral triad	27%	70%
Reported increased interest in Abai	-	84% (self-reported)

These results show a significant improvement in moral concept recognition and cultural understanding.

The increase is consistent with prior findings on corpus-based moral vocabulary instruction [12,14].

Qualitative findings (open-ended responses):

Students described the concepts of *namys* and *iman* as "more than just words" and "guides for living as a Kazakh". One student noted: "Before, I memorized Abai's words for exams. Now I understand what he meant about being human (*adam bolu*)". University students provided more abstract reflections, such as: "Abai's triad teaches harmony of ethics, reason, and will - a foundation for personal integrity".

4.3. Comparative Reflections: School vs. University Learners

The analysis showed differences in cognitive and emotional responses based on education level (**Table 2**):

Table 2. Differences in cognitive and emotional responses.

Category	Secondary Students	University Students
Engagement with corpus	High, especially when visual/concrete	Moderate, but more autonomous
Interpretation of concepts	Emotionally expressive, often personal	Analytical, linked to philosophical frameworks
Vocabulary application	Developing: tendency to paraphrase	Precise and culturally contextualized
Group interaction	Collaborative, oral discussion-based	Reflective, often individual or written

4.4. Task-Specific Student Engagement and Linguistic Insights

During the 60-minute task session, students engaged with four main corpus-based activities. The effectiveness of each task was evaluated through classroom observations, collected outputs (concept maps, reflections, answers), and post-task surveys.

a. Concordance Analysis Task (iman, namys)

Students were presented with concordance lines containing target keywords. For example: *…imandy adam-shyndyqty suip, adal omir suredi…; …namysty el-eshqashan qor bolmas…*

Most secondary students initially paraphrased the meanings ("imandy -good man", "honor-associated with shame"), but after guided corpus observation, 87% of participants recognized these as moral action terms, not just abstract values. One secondary student wrote: *"Before, namys felt like just being hot-tempered or prideful. Now I understand that for Abai, namys means responsibility before the nation and one's own conscience"*. University students approached the concordance with more abstract interpretation: *"Abai frames iman not only as religious faith but as a deep internal ethical compass guiding human behavior"*.

b. Concept Mapping: Tolyq Adam Model

Students constructed visual concept maps around *aqyl-qairat-zhürek*. Examples of mapped relationships: *aqyl* (intellect) → *bilim* (knowledge), *tusinu* (understand), *tandau* (choice); qairat → *areket* (action), *tabandylyq* (perseverance), *kush* (strength); *zhurek* → *niet* (intention), *makhabbat* (love), *adaldyq* (honesty).

Some groups connected Abai's triad with modern dilemmas (e.g., *"Aqyl bar, zhurek zhoq adamdar-tehnologiya damytyp, adamdy umytady"* (People who have intellect but lack heart develop technology yet forget about humanity). These reflections illustrated students' ability to transfer conceptual understanding to contemporary life contexts.

One Grade 10 student explained: *"In my opinion, the tolyq adam is not just a learned person, but someone with a heart who loves their nation. That's what Abai is talking about"*.

University-level responses were more philosophical: *"The triadic model symbolizes equilibrium. Aqyl alone leads to cold logic; zhurek without aqyl risks naïveté. A balanced individual is a socially responsible citizen"*.

c. Keyword Frequency Interpretation

Given a table of keyword frequencies (e.g., adam—152, aqyl—88, iman—41), students interpreted lexical emphasis in Abai's works. Most recognized *adam* as the most central concept.

A secondary student's note: *"Abai connects everything to the idea of 'being human.' It seems to be the ultimate goal"*. University students linked frequency with discourse framing: *"High recurrence of adam, aqyl, iman suggests Abai constructs ethical identity through performative language — actions, not mere labels"*. Some also noticed low-frequency but value-rich terms like *qanagat* and proposed them for further exploration.

d. Reflective Writing: Abai and Me

Students' reflective pieces revealed an emotional and cognitive shift. Common themes included: "Before, I used to just memorize the words, but now I understand their meaning"; "It felt like Abai's words were directed at me personally"; "What Abai wrote is still relevant in today's world".

One powerful student reflection: "Faith, intellect, and honor—I now believe these should become the guiding principles of my own life as well". This task proved particularly effective in bridging abstract philosophical concepts and lived experience.

4.5. Observational Insights and Group Dynamics

- Facilitators recorded the following during classroom sessions:
- Secondary students responded well to **visual scaffolding** (tables, diagrams), and benefited from **peer collaboration**.
- University students displayed **independent inquiry**, including self-directed searches within the digital corpus.
- Emotional expressions (*zhurekpen tusindim* (I understood with my heart), Abai menimen soileskendei (It felt like Abai was speaking to me) were more frequent among school learners.
- University students were more likely to **cross-reference** Abai's prose and poetry when justifying interpretations.

This difference mirrors Vygotsky's theory that scaffolding and social dialogue support conceptual development, especially among younger learners [25].

4.6. Summary of Pedagogical Outcomes

The corpus-based approach enabled context-rich understanding of national values, fostered identity reflection through language, increased student autonomy in constructing meaning, and deepened emotional connection to heritage texts.

Quantitative improvements (pre-/post-survey) were reinforced by qualitative depth, suggesting that corpus methods serve not only cognitive but also affective learning objectives.

The Abai Digital Corpus proved to be a highly effective tool for engaging students in moral reasoning and national identity reflection. The pilot study confirmed that:

- Students developed deeper conceptual understanding of Abai's core values;
- Corpus tasks fostered both emotional and analytical engagement;
- Learning outcomes improved measurably within a short intervention period;
- The approach can be scaled and adapted for different age groups and educational settings.

These findings support the broader scholarly consensus that corpus-informed pedagogy enhances identity-oriented learning [13,14] and affirms the specific relevance of Abai's moral framework in contemporary Kazakh education [7,8].

5. Discussion

The findings of this study demonstrate that the Abai Digital Corpus has significant potential to support the development of national identity among students through value-oriented, language-based education. This section interprets the results in relation to prior scholarly work in the fields of corpus linguistics, moral and national education, and Abai studies, and outlines their pedagogical implications.

The study confirms earlier claims that corpus-informed tasks can deepen learner engagement, critical reflection, and linguistic awareness. As Braun [12] highlights, authentic corpora create space for students to observe real language in use, thereby allowing them to interpret meaning in ways that are contextually rich and learner-centered. In the present study, students used corpus tools to trace the usage of morally significant lexicon in Abai's texts (*namys, aqyl, iman*, etc.), moving beyond memorization toward active exploration.

This reflects the data-driven learning (DDL) principles advanced by Boulton [13], who argues that corpus tasks encourage analytical thinking and increase motivation by positioning learners as discoverers of meaning. In our study, students reported that interacting with concordance lines and semantic clusters made abstract concepts like *iman* and *qanagat* more tangible and personally meaningful.

Moreover, Gablasova, Brezina, and McEnery [14] contend that corpus work supports not only linguistic competence but also the development of learner identity, especially when materials are culturally grounded. The corpus-based design of this study enabled students to internalize Kazakh national values by engaging directly with their linguistic representations.

Abai's model of *tolyq adam* (the complete person), based on the integration of *aqyl* (intellect), *qairat* (will), and *zhurek* (heart), offers a unique ethical and philosophical basis for national education [6]. This triadic concept mirrors the humanist and Sufi traditions, placing morality at the center of identity. Through corpus-based analysis, students not only identified these key concepts but were also able to interpret their interrelationships, thus engaging with Abai's philosophy in a deeper way.

M.O. Auezov [7] and Zhubanov [8] have both emphasized that Abai's writings were not only poetic but also pedagogical in nature, meant to educate and elevate the moral consciousness of the Kazakh people. However, as Ilyasova & Tazhibayeva [9] note, traditional methods of teaching Abai in schools often reduce his work to memorization and formalism. The corpus approach applied in this study directly addresses this limitation by promoting interpretive, reflective, and dialogical learning.

The success of secondary students—many of whom demonstrated emotional and ethical reflection in their re-

sponses—also validates Vygotsky's [25] sociocultural theory: that learning is most effective when it involves interaction with culturally significant symbols and concepts.

The relationship between language and identity is central to sociolinguistic theory [1,2]. Language is not merely a tool of communication but a carrier of collective memory and cultural values. In this study, by analyzing how concepts like *qazaq*, *namys*, and *iman* appear and function in Abai's discourse, students gained access to a culturally embedded worldview.

This aligns with Banks's [3] argument that identity-focused education must offer students opportunities to interpret and internalize culturally specific content through active inquiry. The corpus tasks used in this study acted as mediators between language and meaning, enabling students to co-construct understandings of Kazakh identity through the lens of Abai's thought.

Furthermore, as Halstead & Taylor [4] emphasize, moral education must be contextually grounded. The findings show that Abai's lexicon-rich in moral and spiritual nuance—can serve as an ideal foundation for such education when delivered through interactive methodologies.

The ability of students from both school and university levels to engage with moral philosophy through linguistic analysis suggests that Abai's works can be taught more effectively through corpus-informed approaches. These approaches are scalable, adaptable, and relevant to Kazakhstan's educational modernization agenda (Rukhani Zhangyru), which emphasizes the integration of national heritage with innovative learning tools.

While the intervention was limited in scope, its results indicate that integrating corpus linguistics into Abaitanu and other humanities courses can promote both linguistic literacy and cultural consciousness. As McEnery & Hardie [10] suggest, corpus methods can democratize access to complex texts by breaking them into analyzable, learner-accessible units.

In this case, the digital corpus not only facilitated lexical learning but also created a space for moral inquiry—where values like dignity, self-knowledge, and responsibility were discussed not in abstract, but as lived concepts rooted in Kazakh cultural and linguistic reality.

Limitations

This study is exploratory in nature, involving a relatively small sample size and short duration. While the results are promising, further research is needed to confirm the long-term impact of corpus-based moral instruction on students' identity formation. Future studies might also compare different genres (e.g., poetry vs. prose) within Abai's corpus, or integrate additional Kazakh literary figures for a broader cultural scope.

6. Conclusion

The findings of this study compellingly demonstrate that the Abai Digital Corpus offers significant pedagogical value in fostering national identity, deepening cultural reflection, and nurturing moral reasoning among Kazakhstani youth. By applying corpus linguistics methods to the poetic and philosophical texts of Abai Kunanbayev, the research has illustrated how key national values-*namys* (honor), *aqyl* (intellect), *iman* (faith), and *qanagat* (contentment)—can be identified, analyzed, and meaningfully internalized by students within dynamic, data-driven learning environments. These values, while deeply rooted in the historical and spiritual fabric of Kazakh culture, gain renewed relevance when reframed through interactive and analytical educational approaches that resonate with digital-native learners.

This pedagogical model aligns with and reinforces the work of scholars such as McEnery & Hardie [10] and Boulton [13], who have underscored the transformative potential of corpus-based education in enhancing learner autonomy, interpretive engagement, and intrinsic motivation. When such methods are applied to texts rich in cultural and ethical significance—like those of Abai—they go beyond linguistic analysis, serving as powerful tools for civic education and identity formation. This corresponds with the broader claims of multicultural education theorists such as Banks [3], who argue that educational content grounded in cultural heritage promotes social cohesion and critical self-awareness.

At the heart of Abai's writings lies a sophisticated moral and conceptual framework, epitomized by the figure of the *tolyq adam* (complete human being). This model, which integrates intellectual pursuit, spiritual discipline, and ethical responsibility, has long been recognized as a

cornerstone of Kazakh philosophical thought [6,7]. However, traditional pedagogical practices—often centered on rote memorization and passive reception—have historically failed to engage students with the deeper moral and philosophical dimensions of Abai's legacy. As Ilyasova & Tazhibayeva [9] note, the didactic transmission of Abai's texts frequently overlooks the interpretive and reflective processes necessary for internalizing his ideas.

This study advocates for a shift toward corpus-informed pedagogy, emphasizing learner-centered tasks such as discourse analysis, conceptual mapping, and reflective interpretation. These methods empower students to uncover the linguistic structures through which moral values are constructed and to develop their own perspectives grounded in a cultural context. The dual focus on language and meaning-making facilitates not only intellectual engagement but also emotional and ethical resonance with national heritage.

Furthermore, the comparative findings between secondary school and university students reveal the adaptability and scalability of this approach across developmental stages. Drawing on Vygotsky's [25] sociocultural theory, the study affirms that learning is most effective when it occurs through active participation in culturally meaningful practices. Engaging with Abai's texts in a corpus environment enables students to co-construct meaning, negotiate values, and position themselves within an evolving national narrative.

In a broader sense, this research contributes to Kazakhstan's ongoing *Rukhani Zhangyru* (Spiritual Modernization) initiative, which aims to bridge tradition and innovation by modernizing education through the integration of digital technologies and national cultural content. The Abai Digital Corpus emerges as a strategic educational resource capable of revitalizing *Abaitanu* (Abai Studies), transforming it from a static subject into a dynamic, interdisciplinary field that unites literature, ethics, linguistics, digital literacy, and identity education.

While the scope of this pilot study was limited, its implications are far-reaching. The results underscore the transformative potential of corpus-based approaches in reimagining value-based education. By weaving together linguistic form, cultural substance, and critical reflection, the Abai Digital Corpus does more than preserve Abai's legacy—it revitalizes it, positioning his thought as a living, evolving force in the moral and intellectual development of Kazakhstan's future generations. This study thus calls for further research, broader implementation, and institutional support for corpus-informed methodologies as a cornerstone of culturally responsive and future-oriented education in Kazakhstan and beyond.

Author Contributions

Conceptualization, G.M. and A.B.; methodology, G.M.; software, D.T.; validation, K.P., A.S.; formal analysis, A.S.; investigation, A.B.; resources, K.P.; data curation, G.M.; writing - original draft preparation, K.P.; writing - review and editing, K.P.; visualization, D.T.; supervision, A.B.; project administration, G.M.; funding acquisition, G.M. All authors have read and agreed to the published version of the manuscript.

Funding

Institutional Review Board Statement

Not applicable

Informed Consent Statement

Informed consent was obtained from all subjects involved in the study.

Data Availability Statement

The data supporting the findings of this study are available from the authors upon request. To access the data, interested parties should contact the corresponding author. Please note that while the data are not publicly accessible, they can be provided for research purposes upon reasonable request. Any restrictions and conditions on the

of the data will be communicated by the authors. This ensures that the data are used appropriately and ethically, in line with the study's objectives and confidentiality requirements.

Conflicts of Interest

The authors declare no conflict of interest.

References

[1] Smith, A.D., 1991. National identity. Reno: University of Nevada Press: Reno, NV, USA. pp. 41–45.

[2] Smagulova, J., 2008. Language Policies of Kazakhization and Their Influence on Language Attitudes and Use. International Journal of Bilingual Education and Bilingualism. 11(3–4), 440–475. DOI: https://doi.org/10.1080/13670050802148798

[3] Banks, J.A., 2008. Diversity, Group Identity, and Citizenship Education in a Global Age. Educational Researcher. 37(3), 129–139. DOI: https://doi.org/10.3102/0013189X08317501

[4] Halstead, J.M., Taylor, M.J., 2000. Learning and Teaching about Values: A review of recent research. Cambridge Journal of Education. 30(2), 169–202. DOI: https://doi.org/10.1080/713657146

[5] Suleimenova, E.D., Madiyeva, G., Kozhakhmetova, A., 2020. Multilingualism in Kazakhstan: Policy and identity. Nazarbayev University Press: Astana, Kazakhstan. pp. 88–90.

[6] Kunanbayev, A., 1993. Words of Edification. Zhazushy: Almaty, Kazakhstan.

[7] Auezov, M., 1995. The Path of Abai. Zhazushy: Almaty, Kazakhstan.

[8] Zhubanov, A., 2019. Abai Studies: Language and Legacy. Aruna: Almaty, Kazakhstan. pp. 78–81. (in Kazakh)

[9] Ilyasova, Z., Tazhibayeva, A., 2020. Pedagogical innovations in teaching Abai's legacy. KazNU Bulletin, Philology Series. 84(3), 55–61.

[10] McEnery, T., Hardie, A., 2011. Corpus Linguistics: Method, Theory and Practice, 1st ed. Cambridge University Press: Cambridge, UK. pp. 90–95. DOI: https://doi.org/10.1017/CBO9780511981395

[11] Baker, P., 2006. Using Corpora in Discourse Analysis. Continuum: London, UK. pp. 22–28. DOI: https://doi.org/10.5040/9781350933996

[12] Braun, S., 2005. From pedagogically relevant corpora to authentic language learning contents. ReCALL. 17(1), 47–64. DOI: https://doi.org/10.1017/S0958344005000510

[13] Boulton, A., 2010. Data-Driven Learning: Taking the Computer Out of the Equation. Language Learning. 60(3), 534–572. DOI: https://doi.org/10.1111/j.1467–9922.2010.00566.x

[14] Gablasova, D., Brezina, V., McEnery, T., 2017. Exploring Learner Language Through Corpora: Comparing and Interpreting Corpus Frequency Information. Language Learning. 67(S1), 130–154. DOI:

[15] Baimyrza, A., Serikbayeva, A., Pirmanova, K., et al., 2024. The Impact of Linguistic Identity on the Perception of Abai's Poetry. Forum for Linguistic Studies. 6(5), 239–246. DOI: https://doi.org/10.30564/fls.v6i5.6879

[16] Hanauer, D.I., 2001. What we know about reading poetry: Theoretical positions and empirical research. In: Schram, D., Steen, G.J. (Eds.). Utrecht Publications in General and Comparative Literature. John Benjamins Publishing Company: Amsterdam, Netherlands. pp. 107–128. DOI: https://doi.org/10.1075/upal.35.08han

[17] Hani, H.A., Ajmal, M.U., Arsalan Haider, 2024. Examining the Significance of Literature in Language Learning: A Linguistics Analysis of Selected Poems for Educational Purposes. Journal of Classroom Action Research. 3(1), 1–11. DOI: https://doi.org/10.52622/jcar.v3i1.188

[18] 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", 26 April 2024. Tashkent, Uzbekistan.

[19] Ayazbekova, K., Zaurbekova, G., Svetlana, T., et al., 2024. Development of a Poetic Corpus Based on Abai's Poems. Eurasian Journal of Philology: Science and Education. 195(3), 35–43. DOI: https://doi.org/10.26577/EJPh.2024.v195.i3.ph04

[20] Akhmed-Zaki, D., Mansurova, M., Madiyeva, G., et al., 2021. Development of the information system for the Kazakh language preprocessing. Cogent Engineering. 8(1), 1896418. DOI: https://doi.org/10.1080/23311916.2021.1896418

[21] 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", 26 April 2024, Tashkent, Uzbekistan.

[22] Kambarbekkyzy, P.K., Ogizbayevna, Z.G., Abilkassymovna, A.K., et al., 2024. Initiatives for

developing the Kazakh poetic corpus based on Abai's poems. In Proceedings of the International Conference: Theoretical Hypotheses and Empirical Results, 28–29 March 2024, Oslo, Norway. Available from: https://ojs.publisher.agency/index.php/THIR/article/view/3223?utm_ (cited 10 May 2025).

[23] Severskaya, O.I., 2015. Sociolinguistic Poetics: Theory Dictated by Practice. Language and Method. pp. 201–211. (in Russian)

[24] Yessyim, G., 2012. Abai Qunanbaiuly. Foliant: Astana, Kazakhtan, pp.56–58. (in Kazakh)

[25] Vygotsky, L.S., 1978. Mind in society: The development of higher psychological processes. Harvard University Press: Cambridge, MA, USA. pp. 108–110.