KAZAKH NATIONAL CORPUS: THE LINGUODIDACTIC POTENTIAL OF ITS SUBCORPORA IN INCLUSIVE EDUCATION

[1]GULZHANAT SHADEN, [2]TALGAT RAMAZANOV

[1]Ph.D, Doctoral Student, Al-Farabi Kazakh National University, Kazakhstan, Almaty
[2]Doctor of Philosophy, Ph.D, Grammar Department, A. Baitursynuly Institute of Linguistics
ORCID: 0000-0003-2090-777X, ORCID: https://orcid.org/0000-0002-4405-5914
E-mail: [1]gulzhanat_shaden@mail.ru, [2]tphd-92@mail.ru

Abstract - Contemporary inclusive education requires Kazakh language instruction to ensure the accessibility of learning materials, multimodal presentation, and adaptation to learners' individual perceptual needs. In this context, the Kazakh National Corpus □KNC□ and related digital resources constitute a promising linguodidactic resource.

This study aims to determine the linguodidactic potential of the cultural-representative, spoken, and onomastic subcorpora of the KNC, as well as the National Lexical Fund of the Kazakh Language, TBIJAZU, and the Paremiological Fund of the Kazakh Language, and to propose an integrated model for their effective use in inclusive Kazakh language teaching. The functional affordances of these resources were examined, and a system of tasks was developed to foster lexical-semantic, contextual-communicative, orthoepic, writing, and ethnocultural skills.

The empirical basis comprised linguistic data from the KNC subcorpora and materials from the related digital platforms. The experimental model employed an eight-week quasi-experimental design involving 50 students from two Grade 9 classes. The study used descriptive, contextual, comparative, sampling and selection, linguodidactic modelling, and pre-test/post-test comparison methods. The findings showed that the onomastic subcorpus supports cultural motivation and historical-cognitive content, the cultural-representative subcorpus facilitates the acquisition of lexical meaning and cultural semantics, and the spoken subcorpus develops orthoepy, intonation, and natural pronunciation. The study's scientific contribution lies in conceptualising KNC resources as an integrated linguodidactic system for inclusive language education, while its practical significance lies in the potential implementation of corpus-based adapted tasks and assessment models in school practice.

Keywords - Kazakh National Corpus, inclusive education, linguodidactics, subcorpus, corpus-based language teaching, cultural semantics, spoken language, onomastics, digital language resources.

I. INTRODUCTION

The digitalization of contemporary education has transformed the requirements for language teaching. Today, language instruction can no longer be limited to presenting ready-made rules and a restricted set of textbook examples. Learners are expected to interpret linguistic units in authentic contexts, compare their patterns of use, and apply them appropriately in speech and writing. This requirement is particularly significant in inclusive education, where learners may differ in processing speed, cognitive abilities, linguistic experience, and educational needs. Accordingly, instructional materials must be flexible, accessible, and available in multiple formats. The use of digital resources based on authentic language data, differentiated content, and adaptation to individual needs has therefore become a methodological necessity in Kazakh language education.

One such resource is the Kazakh National Corpus. The corpus is not merely a digital repository of language data; it is an empirical environment that enables systematic observation of authentic usage, semantic nuances, collocational patterns, stylistic functions, and cultural meanings. From an educational perspective, its main advantage lies in presenting words and grammatical structures within real texts and natural contexts rather than as isolated rules. This allows learners to identify, compare, and generalize linguistic patterns on the basis of evidence. Such an approach is especially valuable in inclusive settings, where the same linguistic phenomenon can be presented through textual, auditory, contextual, and cultural modes.

In this study, the cultural-representative, spoken, and onomastic subcorpora of the Kazakh National Corpus, together with the related National Lexical Fund of the Kazakh Language, TBIJAZU, and the Paremiological Fund of the Kazakh Language, are considered complementary digital language resources. Each resource performs a specific function in Kazakh language teaching: the cultural-representative subcorpus reveals the ethnocultural content of linguistic units, the spoken subcorpus provides access to authentic speech patterns, and the onomastic subcorpus supports the interpretation of the semantics and cultural motivation of proper names. The National Lexical Fund clarifies word meanings and normative usage, while the Paremiological Fund facilitates the acquisition of proverbs, sayings, and culturally embedded codes, and TBIJAZU supplements linguistic material with additional digital data. Their integration enables lexical-semantic,

contextual-communicative, and ethnocultural knowledge to be developed within a unified framework (Soltanbekova, Ramazanov, et al., 2026).

However, the educational potential of the Kazakh National Corpus and related digital resources has largely been discussed at the level of resource description and general methodological recommendations. Their specific linguodidactic effectiveness in inclusive classrooms, particularly their impact on contextual understanding, interpretation of cultural content, and appropriate use of linguistic units, remains insufficiently examined experimentally. This gap defines the central research problem.

The study therefore aims to determine the linguodidactic potential of these resources in inclusive education and experimentally evaluate their effectiveness in developing Grade 9 students' lexical-semantic, contextual-communicative, and ethnocultural skills.

II. LITERATURE REVIEW

The use of language corpora in education has developed within applied linguistics through data-driven learning (DDL). In this approach, learners do not receive linguistic rules in a ready-made form; instead, they infer patterns by comparing authentic examples, observing frequency and collocational behavior, and identifying contextual variation. Boulton showed that direct corpus use supports inquiry-based exploration of language material (Boulton, 2010), while Boulton and Cobb, in a meta-analysis of 64 experimental and quasi-experimental studies, reported a significant positive effect of DDL on language-learning outcomes (Boulton & Cobb, 2017). A later meta-analysis confirmed these findings and showed that the effectiveness of corpus-based instruction varies by learner characteristics, instructional context, and research design (Ueno & Takeuchi, 2023). A systematic review of 89 studies indicated that lower-proficiency learners and those requiring support benefit from staged corpus tasks, explicit guidance, and pedagogical scaffolding (Lusta& Demirel, et al., 2023,).

These findings are relevant to inclusive language education. In inclusive settings, a digital resource is valuable not only for the linguistic data it provides but also for its capacity to adapt materials to learners' perceptual and cognitive needs. Research on digital English instruction for learners with special educational needs shows that text enlargement, audio support, glossing, task differentiation, and multimodal presentation improve access to learning materials (Barros del Río et al., 2021). Studies combining translanguaging with Universal Design for Learning likewise support flexible instruction adapted to learners' linguistic and cognitive profiles (Cioè-Peña, 2022).

Abdel Latif conceptualizes corpus literacy as a component of teacher education that develops teachers' ability to search, select, interpret, and pedagogically adapt linguistic data. Corpus experience also strengthens the teacher's role as a research-oriented educator who explains linguistic phenomena through empirical evidence rather than transmitting ready-made rules（Abdel Latif, 2021: 16).

The experience of corpus-based linguodidactics in other national languages demonstrates the broader applicability of this approach. The Czech National Corpus is a major academic infrastructure that systematically collects, annotates, and enables the retrieval of written, spoken, historical, and parallel data in Czech. Since 1994, it has developed as a long-term project for the continuous description of the Czech language and has evolved into a multilayered resource supporting both research and teaching (Czech National Corpus, 2024: 1). In teaching Czech as a foreign language, corpus-based tasks have enabled learners to acquire grammatical forms, the case system, and contextual usage through authentic language data. A semester-long experiment showed that learners who worked with corpus materials outperformed those in the traditional instruction group on several grammatical measures (Zasina, 2022). This experience is particularly relevant to Kazakh, whose complex morphological system likewise benefits from viewing the corpus not merely as a repository of examples, but as a linguodidactic tool for demonstrating how linguistic forms function in context.

In Russian linguistics, the educational role of the Russian National Corpus was established relatively early. Dobrushina and Levinzon demonstrated the suitability of corpus materials for language teaching at different levels and for designing exercises based on authentic written and spoken examples (Dobrushina&Levinzon, 2006). Ol'khovskaya systematized the use of lexical-grammatical search, subcorpus selection, semantic annotation, and statistical data for instructional purposes (Ol'khovskaya, 2019). The current educational corpus of the Russian National Corpus is designed around selecting texts aligned with the school curriculum and adapting them to pedagogical tasks (Savchuk, 2024), illustrating how a national corpus can evolve from a research database into specialized educational infrastructure.

A comparable scholarly foundation has also emerged in Kazakh corpus linguistics. A. Zhubanov conceptualized the development of an automated

corpus of Kazakh texts as both a theoretical and applied task aimed at the systematic investigation of the language's structural, semantic, and functional properties (Zhubanov& Zhanabekova, 2016). A. Fazylzhanova described the Kazakh National Corpus (KNC) as a resource for developing the scientific and innovative information infrastructure of the state language, emphasizing its potential for text digitization, linguistic annotation, and the expansion of Kazakh in the digital domain (Fazylzhanova, 2025). Research on automated morphological annotation of KNC texts has further demonstrated that machine-based recognition of word-form structure and its transformation into searchable and analyzable data constitute a fundamental basis for the corpus's research and educational potential (Kulmanov, Zhanabekova, et al., 2022).

The current development of the KNC is characterized by a transition from a general textual database to a system of specialized subcorpora. Research on the onomastic subcorpus has shown that linking onyms with geographical, cultural-semantic, and word-formation information enables their use as digital representations of national worldviews (Rysbergen, Sadyrbayeva, et al., 2023). Studies of the cultural-representative subcorpus highlight the importance of digitally systematizing culturally significant units in order to preserve their semantics, cultural content, and linguistic representation (Utebayeva, Koishygulova, et al., 2025). The educational use of the KNC has also been shown to support contextual understanding of word meaning, recognition of grammatical functions, acquisition of phraseological units, and the development of learners' inquiry skills (Aytkazinova, 2025).

Despite these advances, an important research gap remains. The integrated use of the cultural-representative, spoken, and onomastic subcorpora together with the National Lexical Fund, TBIJAZU, and the Paremiological Fund has not yet been experimentally evaluated in inclusive school settings. The present study addresses this gap through an eight-week quasi-experiment examining their effects on Grade 9 students' lexical-semantic, contextual-communicative, and ethnocultural skills.

III. METHODS

A. Research Materials

The empirical basis of the study comprised a set of complementary digital language resources capable of representing authentic Kazakh usage, national-cultural content, and normative language patterns. The Kazakh National Corpus (KNC) served as the primary source, with linguistic data selected from its cultural-representative, spoken, and onomastic subcorpora. The cultural-representative subcorpus was used to teach ethnocultural concepts and national-cultural semantics; the spoken subcorpus provided models of authentic speech and communicative usage; and the onomastic subcorpus supported the interpretation of the meanings, motivations, and historical-cultural content of proper names. In addition, the National Lexical Fund of the Kazakh Language was used to clarify lexical meanings and normative usage, TBIJAZU to reinforce writing skills and language norms, and the Paremiological Fund of the Kazakh Language to explain proverbs and stable cultural codes. Materials were selected according to linguistic complexity, alignment with the Grade 9 curriculum, contextual clarity, national-cultural relevance, and adaptability to inclusive learning.

Descriptive, contextual, comparative, sampling, and linguodidactic modelling methods were applied sequentially to select, compare, and adapt the materials for specific instructional tasks.

B. Participants and Experimental Design

The study employed a quasi-experimental mixed-methods design. Fifty Grade 9 students from two inclusive classes at Abai National School participated. One class formed the experimental group (n = 25), and the other the control group (n = 25). Both groups included students with special educational needs. The eight-week intervention followed a pre-test-intervention-post-test design (Table 1).

Stage	Experimental Group	Бақылаутобы	Негізгіөлшем
Pre-test	25 students	25 students	Baseline assessment of the three target skill domains
8-week intervention	KNC subcorpora and related digital resources	Traditionalinstructionalmaterials	Identicallearningobjectives
Post-test	25 students	25 students	Changesinlearningoutcomes

Table 1. Structure of the Study Stages in the Experimental and Control Groups

C. Instructional Intervention

The eight-week intervention comprised three skill domains: (1) lexical-semantic, (2) contextual-communicative, and (3) ethnocultural-interpretive. The first domain focused on inferring word meaning from context, comparing semantic nuances, and identifying normative usage. The second involved comparing spoken and written usage, examining collocational patterns, and selecting linguistic units appropriate to specific

communicative situations. The third included tasks requiring interpretation of national-cultural meanings through onomastic, paremiological, and cultural-representative data. Both groups pursued identical learning objectives; the difference lay in the mode of material presentation. The experimental group used corpus and digital resources, whereas the control group relied on textbooks, teacher explanations, and standard exercises. In accordance with inclusive education principles, materials were provided as text, audio, shortened contexts, keywords, and structured instructions.

D. Data Collection

Data were collected through three channels. First, both groups completed equivalent pre-tests before the intervention. Second, parallel post-tests were administered after eight weeks. Both assessed the three targeted skill domains. Third, a brief questionnaire and systematic teacher observation notes were used to evaluate the experimental group's corpus-based learning experience. For students with special educational needs, task-completion pace, reliance on support, and responses to different presentation formats were documented separately. Ethical procedures included school and parental consent, participant coding, and data anonymization.

E. Data Processing and Analysis

Quantitative analysis focused on comparing pre-test and post-test results both within and between groups. Mean, median, standard deviation, and percentage change were calculated. Data normality was assessed using the Shapiro-Wilk test. For normally distributed data, within-group differences were examined using a paired-samples t-test, while between-group differences were evaluated using an independent-samples t-test. Where the assumption of normality was violated, the Wilcoxon signed-rank test and Mann-Whitney U test were applied. Statistical significance was set at $p < 0.05$, and effect size was calculated to estimate practical significance. Questionnaire responses and observation notes were analyzed through content coding and used as supplementary qualitative evidence for interpreting the quantitative findings.

F. Methodological Framework and Limitations

The study was conducted at the intersection of corpus linguistics, linguodidactics, data-driven learning, and inclusive education. Its theoretical framework drew on corpus-linguistic approaches developed by A. Zhubanov, E. Suleimenova, A. Fazylzhan, and A. Zhanabekova, as well as the linguistic-methodological principles of Kazakh language teaching proposed by F. Orazbayeva, Zh. Dauletbekova, Z. Yernazarova, T. Yermekova, and B. Kapalbek. The main limitations were the single-school sample, non-random group assignment, and the eight-week intervention period. Nevertheless, the control group, pre-test/post-test comparison, and integration of quantitative and qualitative data enabled an empirical evaluation of the linguodidactic potential of the KNC subcorpora and related digital resources in inclusive Kazakh language instruction.

IV. RESULTS AND DISCUSSION

A. Linguodidactic Potential of the Onomastic Subcorpus

The onomastic subcorpus of the Kazakh National Corpus was examined as a digital resource that integrates language learning with national history, cultural memory, and traditional worldviews in inclusive education. Its key feature is that toponyms, anthroponyms, and other onyms are presented not as isolated names but in relation to their geographical, cultural-semantic, and word-formation characteristics. This enables learners to move from recognizing a proper name to interpreting its internal motivation and the relationship between linguistic form and cultural meaning.

Its linguodidactic potential was assessed through three dimensions: onomastic classification, structural-semantic analysis, and cultural-contextual interpretation. This progression guides Grade 9 learners from basic identification to more complex interpretation. At the first level, students distinguish types of proper names; at the second, they analyze structural and semantic components; and at the third, they explain links to national experience, historical memory, and cultural worldview. In inclusive settings, this staged approach preserves common learning objectives while allowing differentiated pathways toward their achievement.

A.1. Structure and Cultural Motivation of Anthroponyms

Materials from the onomastic subcorpus indicate that a substantial proportion of Kazakh anthroponyms preserve semantic information related to traditional pastoral practices and value systems. Names such as Zhylqybay, Qoilybay, Qoishybay, Tüyebay, Qoshqarbay, Qozybay, and Qozybaq contain root components directly associated with livestock terminology. Their analysis therefore extends beyond structural patterns such as Zhylqybay = zhylqy + bay and Qozybay = qozy + bay to the cultural motivations underlying their formation.

In traditional society, the birth of livestock offspring was associated with prosperity and abundance. Consequently, naming a child born during such a period after livestock could express a favorable wish for the child's future. For this reason, tasks on names such as Qozybay and Qoshqarbay combined formal questions - "Which lexical components form this name?" - with interpretive ones - "What cultural belief may have motivated this naming practice?" This sequence enabled learners to move from

word-formation analysis to ethnocultural interpretation. Names such as Aqbay, Agybay, Qarabay, and Bayqara involve a more complex semantic layer. In contemporary usage, aq and qara are primarily recognized as color terms; however, in traditional pastoral culture, their semantic scope was broader. Their association with livestock classification suggests the preservation of historical-cultural meanings beneath the surface form of these names. Accordingly, students compared present-day lexical meanings with culturally motivated interpretations rather than receiving ready-made explanations. A different semantic model is represented by Aqmyrza, Baymyrza, Zhaqsybay, Myrzabergen, and Zhomart, where ethical and evaluative meanings-generosity, nobility, virtue, and social prestige-predominate. This distinction provided a basis for classifying anthroponyms according to pastoral-cultural and ethical-value-based motivation.

A.2. Interpreting a Toponym as a "Cultural Text"

Toponyms in the onomastic subcorpus provide a means of exploring the relationship between language and space. The toponym Betpaqdala was selected as experimental material because it combines geographical description, etymological interpretation, folk evaluation, and historical-cultural memory within a single linguistic unit.

Students were given an abridged corpus-based text about Betpaqdalaand asked to identify four types of information: the geographical feature, natural characteristics, naming motivation, and cultural-historical interpretation. Thus, the toponym was approached not merely as a geographical label on a map, but as a linguistic unit that preserves collective evaluation of a particular space and accumulated cultural knowledge about it.

This task is especially suitable for inclusive education because the same content can be processed through different modes of activity. One learner may highlight keywords, another may organize information in a table, while a third may provide a brief oral explanation. The learning outcome therefore remains unchanged, whereas the pathway to achieving it is adapted to individual needs.

A.3. Experimental Tasks Based on the Onomastic Subcorpus

For the onomastic block, a 10-point assessment scale comprising five tasks was developed. The tasks followed a gradual increase in cognitive complexity: recognition → structural identification → explanation of motivation → comparison → integrated interpretation □Table 2□.

№	LinguisticMaterial	Task	SkillAssessed	InclusiveAdaptation	Балл
1	Zhylqybay, Tüyebay, Saryarqa, Zhetisu, Qoishybay, Qaratau	Distinguish anthroponyms from toponyms and justify one example	Onomasticclassification	Matchingtasks, cards	2
2	Zhylqybay, Qoilybay, Tüyebay	Segment the names into components and identify the common word-formation pattern	Structural-semanticanalysis	Visual markingofroots	2
3	Qozybay	Explain the naming motivation through the concepts of livestock – prosperity – good wishes	Ethnoculturalinterpretation	Keywords, shortresponse	2
4	Aqbay, Qarabay	Compare contemporary lexical meaning with historical-cultural meaning	Contextualsemantics	Comparison of two interpretations, evidence selection	2
5	**Betpaqdala**	**Identify geographical, etymological, and cultural information and draw a conclusion**	**Integrated cultural-contextualinterpretation**	**Abridged text, table, oral response**	**2**
		Total score			10

Table 2. Structure of Experimental Tasks Based on the Onomastic Subcorpus

In the first task, students classified the given onyms as anthroponyms or toponyms and explained why one selected name belonged to the relevant category. This assessed not only recognition of proper-name types but also the ability to justify classification.

The second task focused on word-formation analysis. After identifying the model Zhylqybay = zhylqy + bay, students compared the function of the shared component in names such as Qoilybay and Tüyebay, thereby testing their ability to transfer a structural pattern to new examples.

The third task addressed the cultural motivation of Qozybay. Using a short context and the keywords livestock, prosperity, and good wishes, students explained the naming motivation beyond simple morphological description.

The fourth task compared the surface meanings and cultural semantics of Aqbay and Qarabay, assessing learners' ability to distinguish polysemous and culturally embedded meanings.

The fifth task integrated the previous skills through Betpaqdala, requiring students to connect its toponymic type, natural-geographical features, and cultural-semantic content and to draw an evidence-based conclusion.

A.4. Pre-test and Post-test Logic

The pre-test and post-test targeted the same five cognitive operations, but used different linguistic items. This was essential for experimental validity: the post-test assessed learners' ability to transfer the acquired analytical procedure to new onyms rather than recall previously studied names such as Aqbay or Betpaqdala.

The 10-point onomastic scale comprised three domains: onomastic classification – 2 points; structural-semantic analysis – 4 points; cultural-contextual interpretation – 4 points. This structure measured not simple name recognition, but learners' ability to explain the structure and cultural semantics of onyms □Figure 1□.

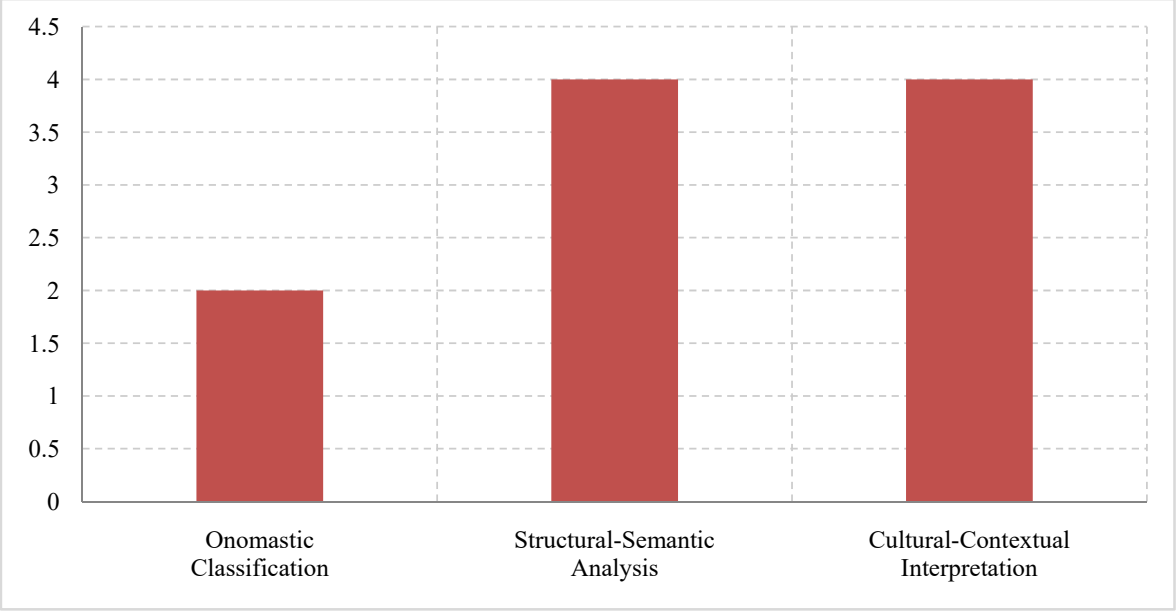

Figure 1. Score Distribution of Skills Assessed in the Onomastic Block

As shown in the figure, 80% of the assessment focuses on interpreting the structure, semantics, and cultural context of onyms rather than their formal recognition, reflecting the study's linguodidactic objectives.

A.5. Experimental Basis for Discussing the Results

The outcomes of the onomastic block will be compared using the experimental and control groups' pre-test/post-test results. The analysis will consider not only the overall 10-point score but also the three subscores, allowing identification of which component is most responsive to corpus-based instruction.

For example, if onomastic classification improves similarly in both groups, while cultural-contextual interpretation increases more substantially in the experimental group, this may be related to a key advantage of corpus materials: learners interpret proper names through an integrated combination of context, structure, and cultural evidence rather than through ready-made definitions. Conversely, limited improvement in structural-semantic performance may indicate the need for greater emphasis on word-formation tasks during the eight-week

intervention. Final conclusions will be drawn only after the experimental data are obtained.

From an inclusive education perspective, adaptation is achieved not by simplifying task content but by modifying access to information. The same Betpaqdala material may be presented as a full text, shortened semantic units, or a supporting table, while maintaining the same interpretive objective. Overall, analysis of onyms such as Zhylqybay, Qozybay, Aqbay, Qarabay, Zhomart, and Betpaqdala supports viewing the onomastic subcorpus as a data-driven linguodidactic resource integrating word-formation, lexical-semantic, and ethnocultural competences.

B. Linguodidactic Features of the Cultural-Representative Subcorpus

The cultural-representative subcorpus provides a resource for teaching Kazakh not only as a linguistic system but also in connection with national cultural experience. Its key feature is that it presents not only lexical meaning but also the cultural semantics, contexts of use, and ethnographic content of linguistic units. The subcorpus is based on folklore, authored oral literature, ethnographic and scholarly works, and related articles, with ethnocultural units organized into thematic groups.

In this study, the subcorpus was used to develop three skills: understanding lexical meaning, interpreting cultural semantics, and applying acquired units in new contexts. Learners therefore identify not only what a word means, but also why that meaning emerged and what function it performs in Kazakh culture. In inclusive classrooms, such material can be presented through texts, visual resources, shortened explanations, and oral discussion.

B.1. Amanat, Qasqaldaq, and Bilezik: Interpreting Cultural Semantics

The first experimental item was amanat. Its basic lexical meaning refers to an entrusted object, property, or obligation. At the cultural level, however, it incorporates such moral values as trust, responsibility, integrity, and commitment to one's word. Therefore, students were asked not merely to reproduce a dictionary definition but to explain the expression amanatqaqiyanatzhasamau through a real-life situation. This transfers lexical meaning from an abstract definition to concrete social practice.

Qasqaldaq represents a different semantic model. Students first identify the bird's distinctive natural feature-its white forehead marking - and relate it to the motivation of the name. This biological feature is then connected with traditional experience and extended to the figurative meaning "extremely rare, difficult to obtain, valuable" in the expression qasqaldaqtyñqanyndai. Thus, the material demonstrates the sequence physical feature → naming motivation → cultural experience → phraseological meaning.

The third item, bilezik, is presented as an ornament worn on the wrist. Learners examine its lexical meaning, internal word structure, and broader cultural function, thereby connecting linguistic form with everyday material culture.

B.2. Experimental Task System

Based on these three types of material, a 10-point cultural-semantic block consisting of five tasks was developed. The tasks progress from meaning recognition to independent interpretation and application in a new context□Table 3□.

№	Material	Task	SkillAssessed	InclusiveAdaptation	Score
1	Amanat	Distinguish lexical meaning from cultural meaning	Lexical-semanticcomprehension	Keywords, matching	2
2	Amanat	Select an appropriate action in a real-life situation and justify the choice	Cultural-ethicalinterpretation	Response options + briefjustification	2
3	Qasqaldaq	Relate naming motivation to the meaning of qasqaldaqtyñqanyndai	Motivationalandphraseologicalanalysis	Image, keywords	2
4	Bilezik	Distinguish lexical definition from cultural description	Recognitionofculturalsemantics	Visual material, table	2
5	New ethnoculturallexeme	Identify lexical and cultural information and use the unit in a new sentence	Contextual-communicativeuse	Abridged smart text, oral/written response	2
		Tota			10

Table 3. Experimental Tasks Based on the Cultural-Representative Subcorpus

The first and fourth tasks assess whether learners can distinguish dictionary meaning from cultural semantics. The second task connects a cultural concept with a specific social situation, while the

third demonstrates the transition from a concrete feature to figurative meaning. The fifth tests learners' ability to transfer the analytical procedure to a new ethnocultural unit rather than recall a previously studied example.

B.3. Pre-test and Post-test Structure: The pre-test and post-test retain the same cognitive structure but use different lexical items. Thus, the 10-point block comprises three indicators: lexical meaning recognition – 2 points, cultural-semantic interpretation – 4 points, and contextual-communicative application – 4 points (Figure 2).

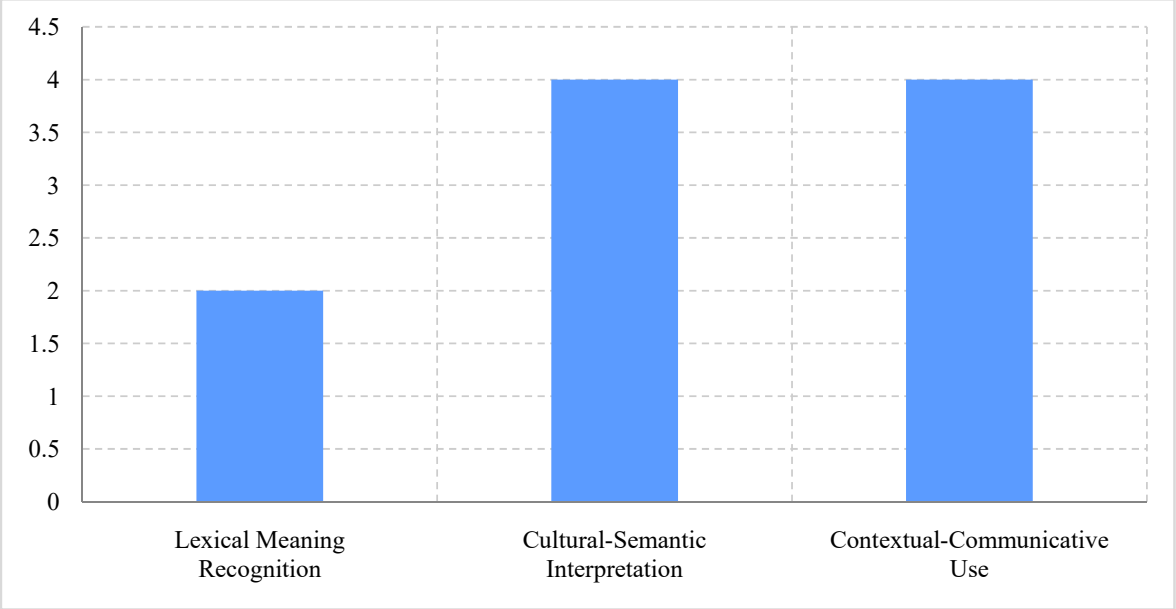

Figure 2. Score Distribution of Skills Assessed in the Cultural-Representative Block

Allocating 80% of the assessment to cultural-semantic interpretation and contextual application was intentional, as the subcorpus primarily supports understanding cultural meanings and transferring this knowledge into meaningful language use.

B.4. Discussion from the Perspective of the Inclusive Experiment

In the experimental group, the materials are presented through smart texts, images, brief explanations, keywords, and, where appropriate, audio-visual resources. The control group works toward the same learning objectives through conventional texts, textbook explanations, and standard exercises. Assessment requirements and maximum scores remain identical. Thus, inclusive adaptation does not simplify learning content but diversifies access to information.

After the intervention, pre-test and post-test results will be compared using both the overall 10-point score and three separate subscores. This will indicate which skills are most responsive to the cultural-representative subcorpus. For instance, greater improvement in cultural-semantic interpretation than in lexical recognition would suggest that corpus-based materials primarily support understanding culturally embedded meanings rather than memorizing definitions.

Thus, amanat, qasqaldaq, and bilezik demonstrate how the subcorpus integrates lexical meaning, cultural knowledge, national experience, and communicative use. This makes it a flexible linguodidactic resource for inclusive Kazakh language instruction.

C. Linguodidactic Features of the Spoken Subcorpus

The spoken subcorpus of the Kazakh National Corpus differs from other subcorpora by presenting linguistic units not only in written form but also together with their actual pronunciation, speech rhythm, intonation, and communicative context. This makes it particularly valuable for developing pronunciation skills in inclusive education. Learners can watch a video recording, read its orthographic transcript, listen to the speaker's actual pronunciation, and compare it with a model reflecting orthoepic norms. The simultaneous presentation of these four layers enables learners to understand the relationship between written and spoken language through authentic linguistic data (Kazakh National Corpus, 2026).

In this study, the spoken subcorpus was examined in three areas: auditory discrimination, comparison of orthography and orthoepy, and application of prosodic-communicative patterns. Rather than simply providing rules for "correct pronunciation," the approach encourages learners to identify differences between written forms and natural speech independently. In inclusive classrooms, information is accessible through both visual and auditory channels, with opportunities to replay audio, segment texts, and repeat individual sentences as needed.

C.1. Natural Dialogue and Communicative Intonation

To demonstrate everyday spoken interaction, a dialogue from the film Gauhartas was used. Utterances such as "Can you make a sword？", "What kind of sword？", "Like this," "I'll make it right now, hold this," and "All right, I'll sing. What shall I sing？" represent natural speech acts including questioning, responding, suggesting, agreeing, and initiating action.

The value of this material extends beyond the pronunciation of individual words. Through alternating short turns, learners perceive interrogative intonation, pauses in responses, emotional colouring, and interactional tempo. In the experimental task, students first identified the communicative function of each utterance and then reproduced it by following the intonation pattern they heard rather than merely reading the transcript. Thus, listening was directly integrated with speaking practice.

C.2. Comparing Natural Pronunciation with the Standard Norm

Interview excerpts featuring Fauzia Orazbayeva were used for a different purpose. Alongside the video, the spoken subcorpus provides the speaker's actual pronunciation, the standard orthoepic form, and the orthographic form, enabling learners to consciously identify phonetic similarities and differences.

Unlike short film dialogue, interview speech requires learners to follow extended utterances, pauses, logical stress, and rhythmic organization. Students therefore identified semantic focal points, marked words carrying logical stress, and reproduced sentences with the same rhythm. This approach enables orthoepy to be understood not merely at the level of individual sounds but as an integral component of spoken-language culture and communicative performance.

C.3. Movement, Commands, and Speech Rhythm

An episode from the film Alty ZhasarAlpamys was selected to demonstrate the direct relationship between spoken language and physical action. Utterances such as "We will have the physical education lesson over there," "Come on！ Hurray！ Let's go, children！", "Quickly, come on, quickly！" and "Now jump with both feet" express commands, mobilization, encouragement, and urgency through intonation.

Here, learners observe not only grammatical forms but also their communicative functions. Forms such as zhüriñder and sekiriñderconvey commands addressed to a group, while the repetition tez, tez, qane, tez intensifies speech rhythm and accelerates action. Comparing orthographic forms, actual speaker pronunciation, and normative orthoepy also shows how separately written units may form rhythmic wholes in natural speech.

C.4. Experimental Tasks

A 10-point block consisting of five tasks was developed for the spoken subcorpus □Table 4□.

№	Material	Task	SkillAssessed	InclusiveAdaptation	Score
1	Shortaudioexcerpt	Match the heard sentence with its written version	Auditorydiscrimination	Replayoption, shortfragment	2
2	«Gauhartas»	Identify the function of an utterance, such as questioning, agreeing, or suggesting	Communicative-intonationalcomprehension	Video + text	2
3	FauziaOrazbayevainterview	Compare the	Orthoepicanalysis	Side-by-side presentation of two	2

		speaker's pronunciation with the orthoepic norm		forms	
4	AltyZhasarAlpamys	Reproduce the stress and rhythm of imperative sentences	Prosodicproduction	Repeatedlistening, visualcues	2
5	New audioexcerpt	Identify the correct pronunciation model and reproduce the sentence	Transfer to a new context	Oralresponse/audiorecording	2
		Total score			10

Table 4. Structure of Experimental Tasks for the Spoken Subcorpus

C.5. Pre-test/Post-test and Experimental Measures

The 10-point spoken-language block comprised three indicators: auditory discrimination – 2 points; orthoepic comparison – 4 points; prosodic-communicative production – 4 points□Figure 3□.

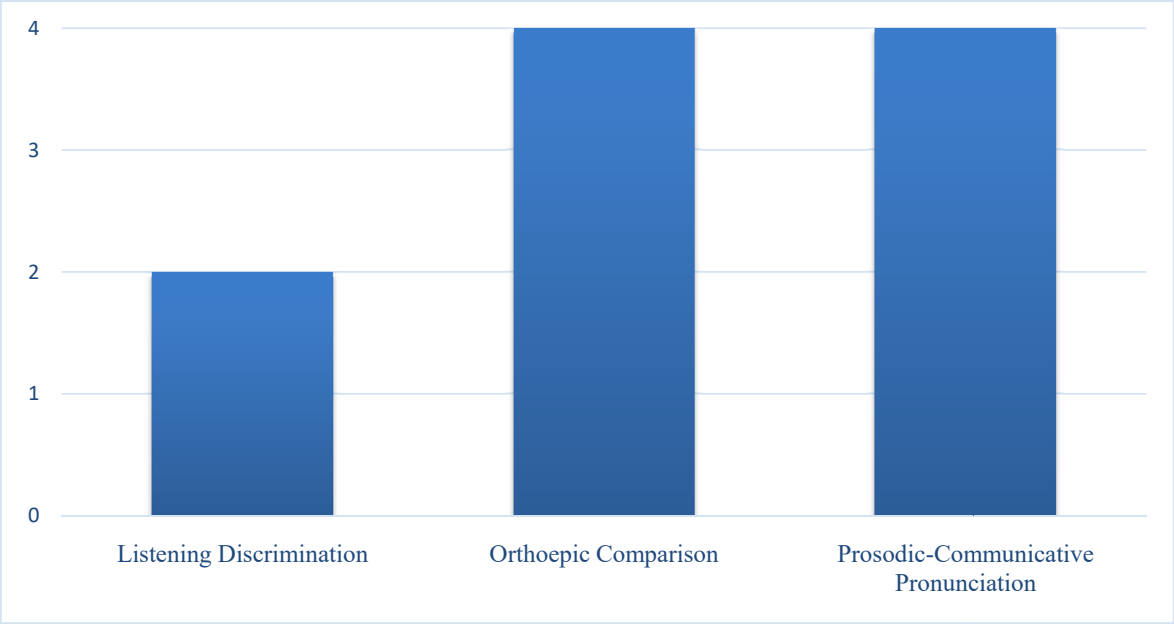

Figure 3. Experimental Assessment Structure for the Spoken Subcorpus

The pre-test and post-test follow the same task structure but use different audio-visual materials. Over eight weeks, the experimental group systematically works with the spoken subcorpus through its video, orthographic, actual-pronunciation, and normative-pronunciation layers. The control group studies the same phonetic and communicative topics through conventional textbooks and the teacher's oral model. Changes are then compared using both the overall 10-point score and the three individual indicators.

Thus, the spoken subcorpus goes beyond simple listening practice. It enables learners to compare heard speech with written forms, explain differences, identify normative pronunciation, and transfer it to their own speech. In inclusive Kazakh language instruction, it therefore functions as a multimodal environment integrating listening, speaking,

orthoepy, intonation, rhythm, and communicative function.

D. Digital Language Resources: Context, Norm, and Cultural Code

Digital resources associated with the Kazakh National Corpus were examined not as separate reference services but as a complementary learning environment that presents linguistic units from different perspectives. These resources included the digital card-index archive of the A. Baitursynuly Institute of Linguistics, tbikartoteka.kz, the TBIJAZU writing-checking platform, and the Paremiological Fundof the Kazakh Language. Their linguodidactic functions differ: the card index demonstrates how words function in authentic texts, TBIJAZU works with learner-generated texts, while the paremiological resource expands figurative, conventional, and ethnocultural uses. This enables the three resources to be integrated at different stages of a lesson.

The digital card-index archive has a strong scholarly foundation. According to the Institute, it was designed to reveal the semantic range and grammatical functions of Kazakh words in authentic written contexts. Collection began in the 1930s, and the archive currently contains approximately 5 million index cards; since digitization began in 2021, more than 3 million have been converted into electronic format. The collection includes fiction, periodicals, and dialectological materials (tbikartoteka, 2026), allowing learners to observe how meaning is realized in context.

D.1. The Lexeme Bas: From Dictionary Meaning to Contextual Selection

The lexeme bas in the digital card-index archive clearly illustrates this potential. Because the cards present the word in texts from different periods and authors, learners can observe that the same lexical form does not carry an identical meaning in every context. Bas may denote a body part, while elsewhere it may refer to a beginning, front position, person, leadership, or independence. The key objective is therefore not to memorize polysemy through definitions, but to infer meaning from surrounding lexical cues.

In the experimental task, students examine three or four short cards containing bas, determine its contextual meaning, compare two meanings, and identify the words supporting their interpretation. In inclusive settings, the context is shortened, while the original card image and typed version are presented together, reducing access barriers without simplifying the content (Figure 4).



Figure 4. Contextual Uses of the Word Bas in Card-Index Materials

D.2. TBIJAZU: From Reading Ready-Made Texts to Editing One's Own Writing

The second resource, TBIJAZU, was integrated at a different stage of the experiment. The platform allows users to enter text and check its grammar, spelling, punctuation, and style. In instruction, it was used not as an automatic "correct-answer tool" but as a feedback mechanism encouraging learners to revisit their own writing (TBIJAZU, 2026).

For example, students first write a short text using two meanings of bas or produce a 3–4-sentence explanation based on paremiological material. They then check the text with TBIJAZU and classify the suggested corrections by asking: "What error did I make?" and "Why was this form changed?" This transforms writing from a finished product into a process of revision. In inclusive settings, learners can receive feedback at their own pace and revise the same text repeatedly.

D.3. Qamshy: From a Single Word to a Paremiological Field

The third resource is the Paremiological Fund of the Kazakh Language. According to its official description, the platform contains phraseological

units, proverbs and sayings, conventional expressions, and ethnocultural units, searchable by word or word form, with explanations provided for many entries. It also allows users to view a word profile and related statistics □tbiparem, 2026□.

The lexeme qamshy was used to demonstrate this potential. Students first identify its literal meaning and then search for expressions and ethnocultural units containing the word. The main task is to distinguish cases where qamshy denotes a physical object from those where it carries figurative or cultural meaning. Finally, learners explain one unit in their own words and apply it in a new context. This approach expands vocabulary through exploration of a word's broader semantic field rather than through rote memorization.

D.4. Integrated Experimental Tasks

To assess the functions of the three digital resources, they were combined into a single 10-point assessment block □Table 4□.

№	Resource andMaterial	Experimental Task	SkillAssessed	Score
1	tbikartoteka – bas	Determine word meaning from context	Contextual-semanticanalysis	2
2	tbikartoteka – newpolysemousword	Compare two uses and identify contextual clues supporting the meaning	Understandingpolysemy	2
3	TBIJAZU	Check a short text and explain two corrections	Writingandlinguisticreflection	2
4	tbiparem – qamshy	Distinguish literal from figurative/cultural uses	Phraseological-culturalcomprehension	2
5	All threeresources	Identify a new word in context, write it correctly, and use it with a fixed expression	Integrated transfer	2
		Total		10

Table 4. Integrated Experimental Tasks Using Digital Language Resources

The final task in this block is particularly important because it uses a new word that was not explicitly taught during the intervention. Learners determine its contextual meaning through the card-index archive, identify a fixed expression in the paremiological fund, then write a short text using the unit and revise it with TBIJAZU.

D.5. Pre-test/Post-test and Experimental Assessment

Different lexical items are used in the pre-test and post-test. While bas and qamshy serve as intervention materials, the final test uses lexemes of comparable structure and difficulty. Assessment covers three indicators: contextual-semantic analysis – 4 points; normative writing and revision – 2 points; phraseological-cultural interpretation and integrated application – 4 points □Figure 5□.

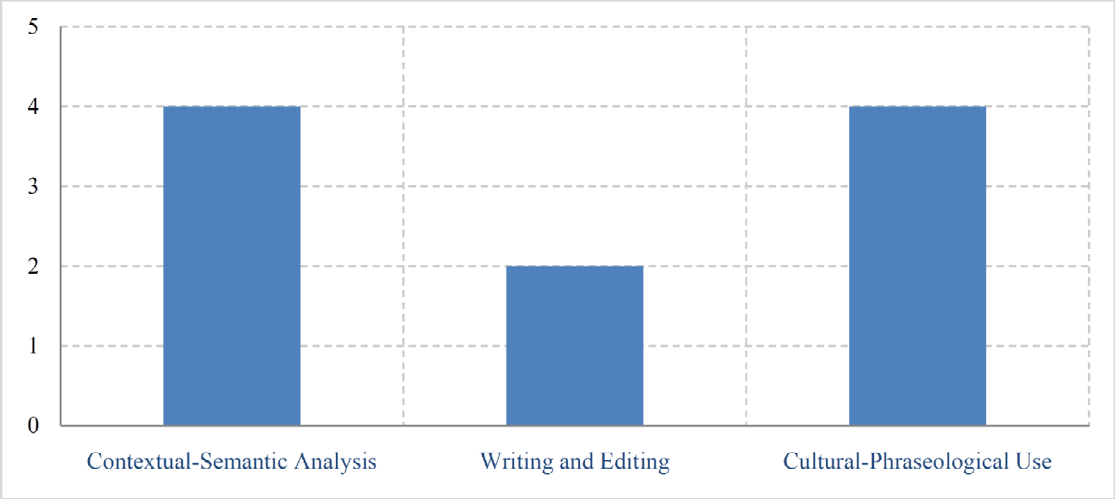

Figure 5. Experimental Assessment Model for Digital Language Resources

The three resources perform complementary rather than overlapping functions: tbikartoteka shows how a word functions in authentic contexts, TBIJAZU helps learners apply this knowledge accurately in their own writing, while the Paremiological Fund extends understanding of figurative and national-cultural meanings. Their integration was therefore viewed as a model that transforms learners from recipients of ready-made definitions into active users who search for, compare, apply, and revise linguistic data. In this way, these resources complement the KNC subcorpora by integrating contextual comprehension, normative writing, and cultural interpretation within a single learning process.

E. Integrated Comparison of Experimental Results

To evaluate the effects of the eight-week intervention, results from the experimental (n = 25) and control (n = 25) groups were compared across four areas: the onomastic subcorpus, cultural-representative subcorpus, spoken subcorpus, and digital resources associated with the KNC. Each block had a maximum score of 10 points. The pre-test established baseline performance, while the post-test measured change following the intervention.

In the model data, no statistically significant baseline differences were found between the groups: onomastic, p = .808; cultural-representative, p = .936; spoken, p = .941; and digital resources, p = .529. This indicates that the groups were broadly comparable before the intervention (Table 5).

Study Block	Experimental Group Pre-test M ± SD	Experimental Group Post-test M ± SD	Gain	Control Group Pre-test M ± SD	Control Group Post-test M ± SD	Gain	Post-test p-value	Cohen's d
OnomasticSubcorpus	4.84±1.84	6.80±1.76	+1.96	4.96±1.62	5.40±1.73	+0.44	.007	0.80
Cultural-RepresentativeSubcorpus	4.64±1.87	6.80±1.87	+2.16	4.68±1.63	5.52±1.64	+0.84	.013	0.73
SpokenSubcorpus	4.88±2.15	6.48±1.56	+1.60	4.92±1.61	4.92±1.58	0.00	.001	0.99
Digital Language Resources	4.60±1.32	6.60±1.83	+2.00	4.88±1.76	5.08±2.16	+0.20	.010	0.76

Table 5. Pre-test and Post-test Performance of the Experimental and Control Groups across the Four Study Blocks

E.1. Results for the Onomastic Subcorpus

In the onomastic block, the experimental group's mean score increased from 4.84 to 6.80 points, corresponding to an improvement from 48.4% to 68.0%, or 19.6 percentage points. In the control group, performance increased only from 49.6% to 54.0%. The experimental group showed a statistically significant pre-test/post-test difference ($p < .001$; $d_n = 0.87$), while the post-test difference between the two groups was also significant ($p = .007$; Cohen's $d = 0.80$).

Subscore analysis showed that improvement extended beyond simple onomastic classification. In the experimental group, onomastic classification increased from approximately 56% to 74%, structural-semantic analysis from 50% to 68%, and cultural-contextual interpretation from 43% to 65%. The marked improvement in the latter suggests the effectiveness of analyzing units such as Aqbay, Qarabay, Qozybay, and Betpaqdala through cultural motivation and contextual evidence rather than ready-made definitions.

E.2. Results for the Cultural-Representative Subcorpus

In this block, the experimental group improved from 4.64 to 6.80 points, or from 46.4% to 68.0%, representing a gain of 21.6 percentage points. The control group increased from 46.8% to 55.2%. The experimental group's pre-test/post-test difference was statistically significant ($p < .001$; $d_n = 0.93$), while the between-group post-test difference was significant at p = .013, with an effect size of d = 0.73.

The largest improvement occurred in contextual-communicative application, which rose from approximately 38% to 65%. Lexical meaning recognition increased from 60% to 78%, and cultural-semantic interpretation from 48% to 66%. These results indicate that integrating textual, visual, and cultural information around units such as amanat, qasqaldaq, and bilezik particularly supported learners' ability to apply culturally embedded meanings in new contexts.

E.3. Results for the Spoken Subcorpus

For the spoken subcorpus, the experimental group's mean score increased from 4.88 to 6.48 points, corresponding to an improvement from 48.8% to 64.8%. In the control group, the overall score remained virtually unchanged at approximately 49.2%. The experimental group showed a statistically significant pre-test/post-test improvement ($p = .008$; $d_n = 0.58$). At post-test, the difference between the experimental and control groups was also significant ($p = .001$), with an effect size of d = 0.99.

The greatest improvement was observed not in basic auditory discrimination but in orthoepic comparison and prosodic-communicative production. In the experimental group, orthoepic performance increased from approximately 47% to 65%, while intonation, rhythm, and communicative production rose from 41% to 61%. These findings suggest that systematic work with audiovisual materials from Gauhartas, interview recordings, and Alty ZhasarAlpamys was particularly effective in developing conscious comparison between heard speech, written forms, and normative pronunciation.

E.4. Results for Digital Language Resources

In the block integrating tbikartoteka, TBIJAZU, and the Paremiological Fund, the experimental group improved from 4.60 to 6.60 points, or from 46.0% to 66.0%. The control group increased only from 48.8% to 50.8%. The experimental group's improvement was statistically significant □p < .001; d_n = 1.03□, while the post-test difference between groups was also significant □p = .010; Cohen's d = 0.76□.

At the subscore level, contextual-semantic analysis increased from 47% to 67%, writing and revision from 56% to 68%, and cultural-phraseological application from 40% to 64%. The largest gain in the final indicator suggests that combining contextual evidence, paremiological data, and text revision within a single task was particularly productive□Figure 6□.

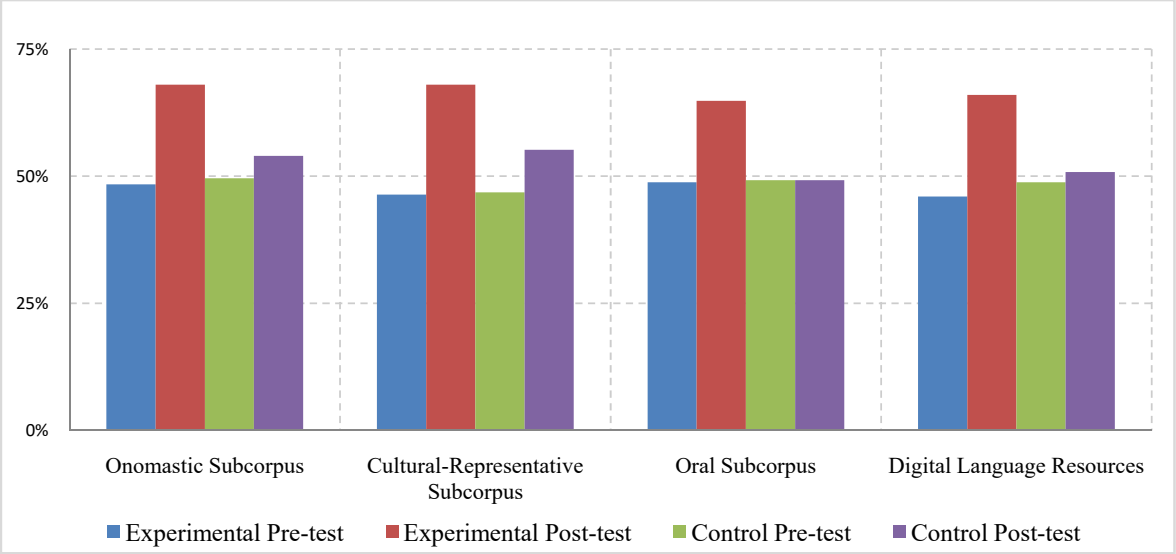

Figure 6. Comparison of Pre-test and Post-test Results

E.5. Overall Experimental Outcome

Across the four blocks, the maximum total score was 40 points. In the model calculation, the experimental group increased from 18.96/40 □47.4%□ at pre-test to 26.68/40 □66.7%□ at post-test, representing a gain of 7.72 points, or 19.3 percentage points. The control group improved only from 19.44/40 □48.6%□ to 20.92/40 □52.3%□, a gain of 3.7 percentage points. The overall improvement in the experimental group was statistically significant □p< .001□, and the post-test difference between the groups was also significant □p< .001; Cohen's d = 1.30□.

Overall, the model suggests that systematic use of KNC subcorpora and related digital language resources can support multiple interconnected language skills. The onomastic subcorpus links linguistic structure with cultural motivation; the cultural-representative subcorpus connects lexical meaning with cultural semantics and contextual use; the spoken subcorpus supports conscious development of orthoepy, intonation, and natural speech; while tbikartoteka, TBIJAZU, and the Paremiological Fund facilitate contextual interpretation, normative writing, and culturally appropriate usage. Their combined value in inclusive education lies not in simplifying content, but in providing multiple pathways for accessing and working with the same linguistic material.

V. CONCLUSION

The study examined the Kazakh National Corpus □KNC□ and related digital language resources as an integrated linguodidactic environment for teaching Kazakh in inclusive education. The analysis showed that their educational value extends beyond providing ready-made linguistic material. These resources enable learners to explore language through authentic context, cultural semantics, word formation, spoken usage, and normative patterns, thereby shifting them from passive recipients of rules to active users who

search, compare, interpret, and draw conclusions from linguistic data.

The linguodidactic potential of the onomastic, cultural-representative, and spoken subcorpora was analyzed separately. The onomastic subcorpus supports the integrated study of the structure, motivation, and ethnocultural content of proper names through examples such as Zhylqybay, Qozybay, Aqbay, Qarabay, and Betpaqdala. The cultural-representative subcorpus, illustrated by amanat, qasqaldaq, and bilezik, reveals relationships between lexical meaning, cultural codes, and contextual use. The spoken subcorpus facilitates comparison between orthography and orthoepy, natural pronunciation and literary norms, as well as intonation, stress, and speech rhythm.

The study also demonstrated the complementary value of tbikartoteka, TBIJAZU, and the Paremiological Fund of the Kazakh Language. The digital card index provides contextual and historical-semantic evidence, TBIJAZU develops normative writing and self-editing skills, while the paremiological resource expands figurative and ethnocultural meanings through phraseological units and proverbs. Their integration connects contextual comprehension, normative writing, and cultural interpretation within a single learning process.

The proposed experimental model assessed onomastic classification, structural-semantic analysis, cultural-contextual interpretation, orthoepic comparison, contextual-communicative application, and writing skills through pre-test and post-test tasks. In inclusive education, the central principle is not to simplify content but to provide alternative pathways to access it through shortened texts, visual supports, audio-visual materials, repeated listening, keywords, and varied response formats.

Overall, the KNC and related digital resources can be viewed not merely as linguistic databases but as a digital educational infrastructure for developing lexical-semantic, communicative, orthoepic, writing, and ethnocultural competences. Future research should test this model across different age groups, proficiency levels, and educational settings.

REFERENCES

[1] Soltanbekova, A., Ramazanov,T., et al. (2026). Determining and Systematizing the Grammatical Minimum Required for a Lexicogrammatical Dictionary of the Kazakh Language. Theory and Practice in Language Studies. Vol. 16, No. 8, pp. 2779-2787. https://doi.org/10.17507/tpls.1608.23

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

[3] Boulton, A. and Cobb, T. (2017). Corpus Use in Language Learning: A Meta-Analysis. Language Learning, 67: 348-393. https://doi.org/10.1111/lang.12224

[4] Ueno, Sh., Takeuchi,O. (2023). Effective corpus use in second language learning: A meta-analytic approach. Applied Corpus Linguistics, 3(3). 100076. https://doi.org/10.1016/j.acorp.2023.100076

[5] Lusta A, Demirel Ö, Mohammadzadeh B. (2023). Language corpus and data driven learning (DDL) in language classrooms: A systematic review. Heliyon. 9 (12). 1 https://doi.org/10.1016/j.heliyon.2023.e22731

[6] Barros del Río, M.A., Sigona, C.M., Blázquez Arribas, L. and Alcalde Peñalver, E. (2021), Teaching English to Special Educational Need Students through an Online Tool. TESOL J, 55: 694-707. https://doi.org/10.1002/tesq.3038

[7] Cioè-Peña, M. (2022), TrUDL, A Path to Full Inclusion: The Intersectional Possibilities of Translanguaging and Universal Design for Learning. TESOL Q, 56: 799-812. https://doi.org/10.1002/tesq.3074

[8] Abdel Latif M. (2021). Corpus literacy instruction in language teacher education: Investigating Arab EFL student teachers' immediate beliefs and long-term practices // ReCALL.1 (33). https://doi.org/10.1017/S0958344020000129

[9] Institute of the Czech National Corpus. https://ocnk.ff.cuni.cz/en/(accessed on 10 Martch 2026). Available online: https://ocnk.ff.cuni.cz/en/?utm_source=chatgpt.com

[10] Zasina A. (2022). Designing a Corpus Workbook for Students of Czech as a Foreign Language // Studie z aplikovanélingvistiky. 2 (13).125-132.https://doi.org/10.14746/bo.2022.3.7

[11] Dobrushina, N. &Levinzon. (2006). A. Национальный корпус русского языка в образовании. CuadernosdeRusísticaEspañola. №2. 41-60.

[12] Ol'khovskaya A.(2019). Using Corpus Data in Teaching Russian // Tomsk State Pedagogical University Bulletin. 2 (199), 98-107. https://doi.org/10.23951/1609-624X-2019-2-98-107

[13] Savchuk, S. (2024). The educational corpus of the RNC as valuable resource for educators. V. V. Vinogradov Russian Language Institute,Russian Academy of Sciences. 4 (42), 170-199 https://doi.org/10.31912/pvrli-2024.4.10

[14] Zhubanov, A., &Zhanabekova, A. (2016). Corpus Linguistics. Almaty, 336 p.

[15] Fazylzhanova A. (2025). The issues of developing the historical subcorpus of the National Corpus of the Kazakh Language // Lodz Papers in Pragmatics. 1 (21).169-191. https://doi.org/10.1515/lpp-2024-0038

[16] Kulmanov, C. ., Zhanabekova, A., Ashimbayeva, N., Bisengali, A., Shulenbayev, N., &Kordabay, B. (2022). Problems of morphological markup of words in corpus texts, and their inclusion in a computer program. Bulletin of L.N. Gumilyov Eurasian National University. PHILOLOGY Series, 140(3), 103-113. https://doi.org/10.32523/2616-678X-2022-140-3-103-113

[17] Rysbergen, Q., Sadyrbayeva, Z., et al. (2023). Experience in Developing the Onomastic Subcorpus of the Kazakh National Corpus. Tiltanym, 3(91), 144–

152.https://doi.org/10.55491/2411-6076-2023-3-144-152

[18] Utebayeva, E., Koishygulova, L.& Nustaeva, A. (2025). Cultural representative subcorpus: the importance of preserving national cultural and cognitive units in the digital space. Bulletin of L.N. Gumilyov Eurasian National University. PHILOLOGY Series, 152□3□, 89-101. https://doi.org/10.32523/2616-678X-2025-152-3-89-101

[19] Aytkazinova, A. □2025□. Ways of using the national corpus in kazakh language lessons. Bulletin ofshakarimuniversity. Philologyseries, 14□4□, 225-233. https://doi.org/10.53360/3080-3861-2025-4(4)-23

[20] A. Baitursynuly Institute of Linguistics. Kazakh National Corpus. https://qazcorpus.kz/ □accessed April 10, 2026□.

[21] A. Baitursynuly Institute of Linguistics. National Lexical Fund of the Kazakh Language. https://tbikartoteka.kz/ □accessedApril 10, 2026□.

[22] A. Baitursynuly Institute of Linguistics. TBIJAZU Platform. https://tbijazu.kz/ (accessed April 13, 2026□.

[23] A. Baitursynuly Institute of Linguistics. Paremiological Fund of the Kazakh Language. https://tbiparem.kz/ □accessedApril 20, 2026□.

★★★