Journal of Culture and Values in Education

https://cultureandvalues.org

E-ISSN: 2590-342X

Volume: 9 Issue: 3 2026

pp. 184-202

Language Interference in Learning Kazakh as a Cognate Language: An Analysis of Turkmen Students' Written Discourse

Alfiya Omurzakova[a], Meruert Satinbekova[a] , Rauza Uskenbayeva[a], Bakytgul Moldagali[b], & Sagira Odanova*[a]

* Corresponding author:
Email: sagira68@mail.ru
a. Department of Kazakh and Russian Languages, Institute of Project Management, Satbayev University (K. Satbayev Kazakh National Technical Research University), Almaty, Kazakhstan.
b. Abai Kazakh National Pedagogical University, Institute of Philology and Multilingual Education, Almaty, Kazakhstan.

Article Info
Received: March 08, 2026
Accepted: August 11, 2026
Published: September 28, 2026

10.46303/jcve.2026.45

How to cite
Omurzakova, A., Satinbekova, M., Uskenbayeva, R., Moldagali, B., & Odanova, S. (2026). Language Interference in Learning Kazakh as a Cognate Language: An Analysis of Turkmen Students' Written Discourse. *Journal of Culture and Values in Education*, *9*(3), 184-202. https://doi.org/10.46303/jcve.2026.45

Copyright license
This is an Open Access article distributed under the terms of the Creative Commons Attribution 4.0 International license.

ABSTRACT
In recent years, the growing number of Turkmen students enrolled in higher education institutions in Kazakhstan has posed new methodological challenges in teaching Kazakh as a cognate language. Although Kazakh and Turkmen are typologically related, differences in their phonetic, morphological, lexical, and graphemic systems result in systematic language interference that undermines the cohesion and coherence of students' written academic discourse. This study aims to identify interference-related errors in the written texts of Turkic-speaking (Turkmen) learners of Kazakh and to examine their impact on textual cohesion and coherence. The study adopts a mixed-methods research design combining qualitative and quantitative approaches. The empirical data consist of summary texts produced by A1-level Turkmen students at a Kazakhstani university. Interference-related errors were identified through Error Analysis, supported by a comparative-typological analysis of Kazakh and Turkmen. Textual cohesion and coherence were examined using discourse models proposed by Halliday, Hasan, and van Dijk. The findings indicate that language interference is systematic and operates across multiple linguistic levels. Morphological interference weakens intersentential grammatical relations, particularly in the use of case, possessive, and personal markers. Graphemic interference, including the mixing of Cyrillic and Latin scripts, further reduces the academic acceptability and readability of students' written texts. Overall, the study demonstrates that interference should be understood not merely as a set of isolated linguistic errors, but as a cognitive–discursive phenomenon that directly affects the quality of written discourse. These findings underscore the pedagogical value of translanguaging and contrastive instructional approaches for teaching Kazakh to Turkic-speaking learners.

KEYWORDS
Language interference; cognate language learning; Kazakh language education; written discourse; cohesion and coherence; translanguaging approach.

INTRODUCTION

The phenomenon of language interference within the Turkic language family represents a unique intersection of genetic relatedness and linguistic divergence, presenting significant challenges for learners in academic settings. As higher education institutions in Kazakhstan experience a substantial influx of international students from neighboring Turkic-speaking nations, particularly Turkmenistan, the traditional models of language instruction designed for non-related learners have proven insufficient (Goodman & Tastanbek, 2020; Karabassova & San Isidro-Smith, 2020). While the genetic proximity of Kazakh and Turkmen facilitates rapid initial acquisition, it also creates a "cognate paradox" in which structural similarities increase the likelihood of negative transfer (Clyne, 2003; Ringbom, 2006). This investigation addresses the critical need to identify the specific variables of interference in the written discourse of Turkmen students, a group currently underrepresented in the linguodidactic models of Kazakhstani universities.

The theoretical foundations of this study are grounded in the Contrastive Analysis Hypothesis (CAH) and Interlanguage Theory. Although CAH originally posited that errors could be predicted solely by comparing L1 and L2 structures, later scholars recognized the learner's language as a separate, rule-governed system–an interlanguage–that reflects evolving hypotheses about the target language (Corder, 1967; Selinker, 1972). In the context of cognate languages, negative transfer is driven by the "priming effect," a psycholinguistic mechanism where the perception of a cognate word in one language automatically activates its mental representation in the other, often leading to triggered code-switching or morphological overgeneralization (Broersma & De Bot, 2006; Dijkstra et al., 1999). Furthermore, this study utilizes the Common Underlying Proficiency (CUP) model, which suggests that academic skills developed in the primary language are transferable to subsequent languages, provided that instructional models acknowledge the learner's existing linguistic repertoire (Cummins, 1981; Karabulut & Kesli Dollar, 2022).

Despite the importance of these phenomena, a specific research gap exists in the analysis of written academic discourse among Turkmen learners. Previous studies have largely focused on phonetic and morphological errors in isolation (Johanson, 2002; Odlin, 1989), whereas the cumulative impact of interference on textual cohesion and coherence remains underexplored (Halliday & Hasan, 1976; van Dijk, 1980). This lack of a systematic description of writing errors in a digital and academic format hinders the development of effective diagnostic and teaching methodologies for Turkic-speaking learners (Amaniyazova, 2020; Imankulova et al., 2025).

The problem addressed in this study is the prevalence of systematic interference-related errors in A1-level Kazakh summaries produced by multilingual learners, which compromise textual accuracy and linguistic naturalness. Accordingly, the study seeks to answer the following research questions: (1) What systematic interference patterns occur in A1-level Kazakh summaries? (2) How does translanguaging-based instruction affect the frequency of these interference patterns**?** To address this, the study adopts a Corpus-Based Error Analysis (CEA)

design, which offers a structured, robust alternative to traditional qualitative observation by examining naturally occurring language data (Granger, 2003; McEnery & Brookes, 2024). The general aims of this research are: (1) to identify the frequency and typology of interference-related errors within a specialized learner corpus of 12,400 words; (2) to distinguish systematic interlanguage patterns from random performance slips; and (3) to evaluate the effectiveness of a translanguaging-based pedagogical model in mitigating these errors (Canagarajah, 2011; García & Li, 2014). By grounding the research in translanguaging, which views the learner's full linguistic repertoire as a resource rather than a hindrance, this study provides a new framework for sustaining the "linguistic naturalness" of Kazakh academic discourse in a multilingual environment (Goodman & Tastanbek, 2020; Tastanbek, 2019).

THEORETICAL FRAMEWORK

Language Interference and Cross-Linguistic Influence

Language interference, or cross-linguistic influence, is defined as the effect of a speaker's native language (L1) on their usage of a second language (L2), which frequently results in errors across pronunciation, grammar, and vocabulary (Weinreich, 1953; Odlin, 1989). This phenomenon commonly occurs in bilingual settings when L1 habits–such as sentence structure, word choice, or morphological paradigms–are transferred to the target language system (Weinreich, 1953). Within the framework of Interlanguage Theory, these deviations are viewed as a dynamic, rule-governed system that draws on L1 knowledge to bridge gaps in L2 proficiency (Selinker, 1972). While traditional Contrastive Analysis initially viewed all transfer as a source of error, modern scholarship recognizes that interaction between two systems is a cognitive strategy in which the learner's prior linguistic experiences inevitably shape their production in the target language (Jarvis & Pavlenko, 2008).

The Cognate Paradox and the Priming Effect

A central variable in this study is the role of cognates–lexical items that share a common etymological origin and exhibit similar forms and meanings across different languages (Levelt, 1989). Cognates function as "cognitive bridges" that facilitate faster word recognition and ease access to the mental lexicon (Cummins, 1981). However, in the acquisition of genetically related languages like Kazakh and Turkmen, learners often encounter a "cognate paradox" – the situation in which the very similarity between cognate forms simultaneously accelerates initial comprehension and raises the risk of negative transfer, because the learner cannot reliably suppress the L1 representation when processing the structurally similar L2 form (Ringbom, 2006). While structural proximity facilitates initial comprehension, it also increases the likelihood of negative transfer because the learner cannot effectively inhibit the L1 when processing highly similar L2 structures (Dijkstra et al., 1999).

This is driven by the "priming effect," a psycholinguistic mechanism where the perception of a cognate in the target language automatically and unconsciously activates its mental representation in the native language (Broersma & De Bot, 2006). Research indicates that this

activation frequently leads to " triggered code-switching" – a phenomenon distinct from code-mixing. Whereas code-mixing refers to the deliberate or habitual blending of elements from two languages within a single utterance for stylistic or identity-marking purposes, triggered code-switching denotes the involuntary shift into the non-target language automatically prompted by the recognition of a cognate form; the cognate acts as a trigger that momentarily activates the L1 system, causing surrounding words to be produced in the L1 rather than the L2 (Broersma & De Bot, 2006). In this study, the term is used in this psycholinguistically grounded sense. In the case of Turkic languages, this interference is most salient in written discourse, where learners overgeneralize similarities in vowel harmony and affixation, leading to systematic errors (Johanson, 2002).

Translanguaging Pedagogy: Definition and Research Examples

The translanguaging approach, first introduced by Cen Williams (1994) and expanded by Ofelia García, reframes the multilingual learner's repertoire as a single, integrated system rather than separate "boxes" for each language (García & Li, 2014; Williams, 1994). It is important to distinguish translanguaging from translation. Translation is a cognitive task in which the speaker renders a fixed message from one language into another, treating the two languages as separate, bounded systems with no overlap. Translanguaging, by contrast, rejects the notion that a multilingual speaker possesses distinct, isolated language systems. Instead, it conceptualizes the speaker's entire linguistic repertoire – across all languages – as a single, flexible, integrated resource from which the speaker draws fluidly and purposefully depending on communicative need and social context (García & Li, 2014). Translanguaging is defined as a pedagogical practice that promotes the use of all a learner's linguistic resources to support learning and meaning-making (García & Kleifgen, 2020).

Research demonstrates several successful examples of this approach:

- **Scaffolding Content:** Qualitative studies in trilingual schools have found that teachers use translanguaging to scaffold complex disciplinary concepts, allowing students to conduct initial processing in their L1 before producing academic output in the L2 (Karabassova & San Isidro-Smith, 2020).
- **Enhancing Writing Achievement:** Experimental studies of EFL learners have shown that students exposed to a translanguaging instructional cycle achieve significant gains in task achievement, lexical and grammatical accuracy, and overall cohesion and coherence (Karabulut & Dollar, 2022).
- **Validating Identity:** By validating the learner's full linguistic identity, translanguaging reduces classroom anxiety and fosters higher engagement in difficult academic tasks (Kaiypova et al., 2025; Wang et al., 2024).

Research Gap: Interference in Written Academic Discourse

Despite the growing body of research on Turkic language contact, this study addresses a specific research gap: the lack of corpus-based, discourse-level analysis of how Turkmen–Kazakh interference affects textual cohesion and coherence in A1-level academic writing. No prior study

has systematically combined Error Analysis across phonetic, morphological, graphemic, and lexical levels with a discourse-structural examination of how those errors accumulate to disrupt the genre conventions of written summaries. The gap is also not just about the limited understanding of how interference affects learners' written discourse (Halliday & Hasan, 1976; Ismagulova et al., 2024; Kuzembayeva et al., 2024). Most existing studies focus on isolated phonetic or morphological errors rather than examining how these errors erode textual integrity (Johanson, 2002). This study specifically intends to explore the genre of written summaries produced by A1-level learners. The summary genre is a critical indicator of academic literacy, as it requires the learner to reorganize and generalize information while maintaining textual cohesion (the formal links between sentences) and coherence (the semantic macrostructure) (van Dijk, 1980; Karlinsky, 1996). By focusing on functional adequacy–specifically comprehensibility and articulation–this study seeks to identify how systematic interference from the Turkmen L1 disrupts the logical development of ideas in Kazakh academic writing (Granger, 2003; Stefanowitsch, 2020).

METHOD

Research Design and Variables

The present study adopts a Corpus-Based Error Analysis Design to provide a comprehensive analysis of language interference (Creswell & Creswell, 2017). This design integrates a quantitative corpus-based approach to identify error frequency with a qualitative discourse analysis to examine the cognitive and discursive causes of these deviations (Brezina, 2018).

As a corpus-based error analysis study, this research does not operate within a variable-centered experimental framework and therefore does not designate independent or dependent variables. Instead, the study is oriented around two focal analytical constructs. The first is linguistic interference, understood as systematic, recurring deviations from Kazakh norms in students' written texts, classified across phonetic, morphological, graphemic, and lexical levels. The second is written discourse quality, operationalized through the constructs of textual cohesion (Halliday & Hasan, 1976) and coherence (van Dijk, 1980), examined qualitatively through discourse-structural analysis of students' written summaries. The central analytical question is how the first construct shapes and disrupts the second.

Participants and Sampling Strategy

The study employed a purposive sampling strategy to select a homogenous group of participants who share the same linguistic profile (McEnery & Brookes, 2024). A total of 62 Turkmen students aged 18–22 enrolled at Satbayev University in Almaty, Kazakhstan. This age range was selected because it represents "digital native" undergraduate learners who are transitioning from the informal Latin-script digital environment of their L1 to the formal Cyrillic-script academic environment of the Kazakh L2 (Tastanbek, 2019; Goodman & Tastanbek, 2020). All 62 participants were learning Kazakh at the A1 proficiency level, ensuring that the errors analyzed

reflect the foundational stages of cognate language acquisition where interference is most salient.

Corpus Construction and Data Collection

The empirical material consisted of a specialized Learner Error Corpus of approximately 12,400 words, derived from 186 authentic written summary texts (three per participant) produced as part of the students' academic coursework. The collection followed natural sampling principles to ensure the data reflected real language use in a classroom setting (Corder, 1975).

The study employed two categories of research instruments aligned with the corpus-error analysis design. The corpus instrument consisted of a standardized error-tagging protocol applied to the full 12,400-word corpus, with errors extracted and classified using PowerConc 1.0 software to generate frequency lists and error-type distributions across linguistic levels (Brezina, 2018). Frequency counts and percentages were computed as descriptive indicators to characterize the distribution and relative salience of interference patterns; these numeric summaries serve an organizational function within the qualitative analysis and do not constitute inferential statistical testing.

The qualitative instrument comprised two components. First, context-rich error samples and discourse-level text segments were systematically analyzed using the frameworks of cohesion (Halliday & Hasan, 1976) and coherence (van Dijk, 1980) to trace how interference patterns disrupt textual integrity. Second, a semi-structured follow-up interview protocol (10 open-ended questions) was administered with a purposively selected sub-sample of 12 participants to elicit reflective accounts of their writing challenges, L1 transfer strategies, and experiences with Kazakh morphological complexity. Interview data were transcribed and subjected to thematic analysis to identify recurring cognitive and discursive patterns not captured by the corpus error tags.

Data Analysis and Trustworthiness

The data analysis proceeded in two integrated phases. The first phase involved systematic corpus error analysis, with frequency counts and percentages used as descriptive organizational tools within the qualitative framework rather than as inferential statistics. The second phase employed qualitative thematic analysis of discourse-level text segments and semi-structured interview data. Because this study follows a corpus-error analysis design, numeric data serve to map and characterize interference patterns; the primary mode of analysis throughout is qualitative and interpretive. To analyze the qualitative data from the interviews, a systematic qualitative coding framework was utilized as the primary analytical tool to identify key themes. *Phase 1: Corpus Error Analysis.* In the first phase, errors in the 12,400-word corpus were processed through the five-step Error Analysis procedure (Corder, 1975; Selinker, 1972): (1) systematic corpus compilation; (2) identification of deviations from Kazakh norms; (3) typological categorization by linguistic level; (4) explanation of the underlying interference mechanisms; and (5) evaluation of functional impact on discourse adequacy. Error tagging employed a standardized coding system based on the FRIDA protocol (Granger, 2003).

Frequency counts and percentages were computed as descriptive indicators to organize and characterize the distribution of interference patterns; these figures support qualitative interpretation and do not constitute inferential statistical analysis (Brezina, 2018). Coding reliability was established by two trained linguists independently classifying a 10% subset of the corpus, yielding a Cohen's kappa of .87, with discrepancies resolved through consensus discussion (McEnery & Brookes, 2024).

Phase 2: Qualitative Analysis and Trustworthiness. Building on the corpus error patterns identified in Phase 1, the second phase employed qualitative thematic analysis of two streams of narrative data (Creswell & Creswell, 2017). First, extended discourse-level text segments from the corpus were analyzed against the cohesion framework (Halliday & Hasan, 1976) and the coherence macrostructure model (van Dijk, 1980), tracing how interference patterns disrupt textual integrity at the sentence-linking and paragraph-organization levels. Second, transcripts from semi-structured follow-up interviews with 12 purposively selected participants were subjected to inductive thematic coding, generating themes that reflect learners' lived experiences of L1–L2 structural conflict in academic writing. Initial open codes were progressively grouped into higher-order themes through constant comparison.

To ensure the trustworthiness of the narrative data process, four procedures were applied: (1) prolonged engagement with the corpus and interview data to reach analytical saturation; (2) member-checking, sharing thematic summaries with three participants for respondent validation; (3) negative case analysis, systematically examining deviant instances to refine emerging themes; and (4) triangulation across corpus error data, questionnaire responses, and interview accounts to strengthen the credibility and confirmability of the interpretive findings. The corpus data were compiled from students' written assignments collected over the duration of the study, while questionnaire and interview data were gathered sequentially to capture participants' perspectives on their writing experiences and perceived linguistic challenges. In addition, all data sources were systematically organized, coded, and cross-referenced throughout the analysis process, enabling the identification of recurring patterns and the verification of interpretations across different forms of evidence.

RESULTS AND DISCUSSION

This section presents the findings in two sequential phases consistent with the study's research design. The first phase reports the quantitative results of the corpus-based error analysis, providing frequency data and typological distributions of interference-related errors across linguistic levels. The second phase presents the qualitative interpretive findings, drawing on discourse-level text analysis and interview data to explain the cognitive and discursive mechanisms underlying the identified patterns. Together, the two phases address the study's four research aims. Quantitative analysis of the 12,400-word learner corpus identified 1,376 interference-related errors. These are organized below by linguistic level, followed by a

qualitative account of their cumulative impact on the cohesion and coherence of students' written discourse.

The overarching aim of this analysis is to determine how systematic L1 (Turkmen) interference shapes the linguistic form and discourse-level quality of A1-level learners' written Kazakh. To achieve this aim, the study seeks to:

(1) Conduct a comparative analysis of the structural similarities and differences between Kazakh and Turkmen.

(2) Systematize interference errors in students' written texts across phonetic and grammatical levels.

(3) Identify patterns of disrupted textual relations based on the genre of written summaries.

(4) Develop methodological recommendations for teaching Kazakh as a cognate language grounded in the findings of the study.

1) Interference-Related Errors at the Phonetic Level

The investigation addressed three primary categories of cognates: Phonological Cognates (similar vowel sounds), Morphological Cognates (shared Turkic affixes), and Lexical Cognates (shared root words). The comparative analysis revealed that the high degree of mutual intelligibility between the Oghuz group (Turkmen) and the Kipchak group (Kazakh) creates a "facilitation trap" in which learners misjudge the structural divergence of shared categories.

Table 1.

Kazakh-Turkmen Cognate Differences and Their Interference Effects

Cognate Variable	Kazakh (L2) Feature	Turkmen (L1) Feature	Resulting Interference
Vowel Cognates	12 Vowels (High Harmony)	9 Vowels (Moderate Harmony)	Violation of positional assimilation
Plural Cognates	12 Allomorphs (-lar, -dar, -tar)	4 Allomorphs (-lar, -ler)	Morphological overgeneralization
Case Cognates	7 Cases (Instrumental exists)	6 Cases (Instrumental missing)	Substitution with postposition 'bilen'
Graphemic Cognates	Cyrillic script (formal)	Latin script (digital/L1)	Digraphia and script-mixing

The initial stage of acquiring Kazakh involves mastering the alphabet and grapheme–phoneme correspondences. At this stage, Turkic-speaking students, including Turkmen learners, frequently produce errors resulting from phonetic mismatches and differences between orthographic systems. The Kazakh alphabet comprises 42 letters, whereas the Turkmen alphabet includes 30 letters. Although some graphemes are shared between the two languages, their phonetic realization and articulatory properties differ, leading to interference-related deviations in written discourse.

For example, the letter **u** in Turkmen is phonetically close to the Kazakh vowel **y**, which results in erroneous spellings such as *suyq* → *suvyk* and *boyau* → *boyaýv*. Additional difficulties

arise in representing sounds that are absent from the Turkmen alphabet or have different graphic realizations.

a) Sounds g ~ ğ

As the letter **ғ** is absent from the Turkmen alphabet, students tend to replace it with **r**. This error is attributable to graphic similarity and articulatory proximity between the two sounds: *ağylşyn → agylşyn, ğalym → galym, jağajai→ jagajai*

b) Phonetic mismatch e ~ i

In Turkmen, the function of the sound **i** is performed by the letter **ý**. Under its influence, students frequently insert an unnecessary **i** before words beginning with **e**, which constitutes a graphemic manifestation of phonetic interference *eden → ieden, esık → iesık, etık →ietık*

c) Sounds i ~ i (ŭ)

Because Turkmen lacks a direct equivalent to the Kazakh sound **i**, students tend to render it as a short **i (ŭ)** or incorrectly combine it with **i/e**. This pattern represents a typical case of interference caused by mismatches in phoneme–grapheme correspondences *biık → beiık, kiım → kim, niet → nıiet, ine → ıine*

d) Sounds k ~ q

The absence of the letter **q** in Turkmen leads to its being replaced by **k** and, in some cases, by **ҳ**. This type of error is associated with the articulatory base of the learners' native language *qalyñ → kalyn, taqta → takta, suyq → suvyk*

e) Sounds ū ~ o, u, ü

Turkmen students frequently confuse the vowels *ū, o, u, ü,* reflecting phonological discrepancies between the vowel systems of Kazakh and Turkmen *ūran → oran, ğūlama → golama, mūz → muz*

f) Sounds u ~ v

The frequent use of the sound **v** in Turkmen (e.g., *suw, okuwçy*) results in its transfer into Kazakh written discourse. This transfer manifests itself in the insertion of an additional **v** following **u** in Kazakh words *boiau → boiav, äuen → även, toqu → toquv*

In other words, when processing a lexeme in the second language (L2), learners automatically map it onto the phonetic representation and graphic model of their first language (L1).

Influence of the Turkmen Alphabet and Graphemic Interference

A distinctive feature of Turkmen students' written communication is the influence of the Latin script used in their native language. The official use of the Latin alphabet in Turkmenistan and the Cyrillic alphabet in Kazakhstan directly affects students' writing practices. It contributes to the emergence of various forms of graphemic interference.

The analysis revealed that when writing in Kazakh using the Cyrillic script, Turkmen students frequently substitute individual Cyrillic letters with Latin graphemes or automatically reproduce orthographic patterns characteristic of the Turkmen language. This phenomenon can be attributed to the digraphic nature of their writing practices and the active use of the Latin

script in digital communication environments, including mobile devices, social networks, and messaging applications.

Accordingly, students' written texts often contain erroneous spellings such as: men → ten *(ìåí → òåí)*, asymdy → asytdy *(àñûìäû → àñûòäû)*, universitet → universitet *(óíèâåðñèòåò → óíèâåðñiòåò)*, it → it (*èò → it)*, dosym → dosyt *(äîñûì → äîñûò)*, minut → minut *(ìèíóò → ìiíóò)*, bazar → bazar (*áàçàð → áàzàð)*. In these examples, Latin graphemes *(a, t, i, z)* and Turkmen orthographic patterns are used alongside Cyrillic letters, resulting in a mixed writing model.

These examples demonstrate the systematic nature of mixing Cyrillic and Latin scripts. Several factors contribute to the emergence of graphemic interference. First, the digital communication factor: students are accustomed to using Turkmen in the Latin script on social networks and messaging applications, and this habit is carried over into academic writing. Second, the visual similarity factor: some Latin letters visually resemble Cyrillic characters but differ in phonetic value *(t, i, z),* leading to erroneous identification. Third, the phonological mismatch factor: discrepancies between spelling–pronunciation correspondences in Turkmen and Kazakh encourage learners to transfer the phonetic–orthographic model of their native language into Kazakh writing.

As a result, graphemic interference gives rise to mixed orthography (mixed script) in the written discourse of Turkic-speaking students. This phenomenon hinders the normalization of written language and directly affects learners' levels of orthographic literacy and overall writing culture.

2) Interference-Related Errors at the Morphological Level

Despite the genetic relatedness between Kazakh and Turkmen, their morphological systems exhibit several substantial structural and functional differences. These similarities and divergences, on the one hand, facilitate the acquisition of certain grammatical forms; on the other hand, they give rise to morphological interference through the mechanism of false analogy.

The analysis revealed that when attaching affixes in Kazakh and Turkmen, students frequently rely on morphemic patterns from their native language and tend to overgeneralize morphological correspondences. As a consequence, systematic errors occur in the use of personal, possessive, plural, and case markers.

a) Interference Related to Personal Endings

Although Kazakh and Turkmen share the same syntactic typology (SOV), the principles governing morphological harmony differ considerably and constitute a major source of grammatical interference. In Turkmen, affixes are primarily governed by vowel harmony (front vs. back vowels), whereas positional consonant assimilation is not consistently realized. In Kazakh, by contrast, the allomorphs of both derivational and inflectional affixes are conditioned not only by syllabic harmony but also by the assimilation of voiced, voiceless, and sonorant consonants, resulting in a wide range of phonetic variants.

As a result of these differences, Turkmen students frequently experience difficulty in selecting appropriate personal forms in Kazakh. Typical interference-related errors include the following patterns: *studentmın → studentmız, ūşqyşpyn → ūşqyşmyn, mūğalımmız → mūğalımbyz, şäkırtpız → şäkırtmyz*. Furthermore, because the form **size** in Turkmen functions both as a polite singular and as a plural form, students often fail to distinguish between the Kazakh forms **sız** and **sızder**, producing forms such as *studentsızder → studentsıñız, oquşysyz → oquşysyñyz, ūşqyşsyñyz → ūşqyşsyz.*

Another persistent interference pattern involves the attachment of an additional plural marker to verb forms in the third-person plural, for example: *olar keldıler, aittylar, bardylar, söiledıler.* This error reflects Turkmen grammatical patterns, in which verb forms accompanying plural subjects are marked by the suffixes *-lar* or *-ler*.

b) Interference Related to Possessive Markers

Although possessive markers in Kazakh and Turkmen are similar in their surface forms, they differ significantly in their usage constraints and vowel harmony patterns. Because the Turkmen form **siz** simultaneously encodes politeness and plurality, students tend to employ plural possessive forms when expressing singular reference in Kazakh.

For instance, in constructions such as *sızdıñ kıtaptaryñyz, däpterleriñız, qalamdaryñyz,* the singular form *sızdıñ* is combined with plural possessive markers. In some cases, this interference also manifests as confusion between the forms *sender* and *sızder*, as in *sızderdıñ esıgıñız, sender studentsıñızder.*

In addition, violations of labial vs. non-labial vowel harmony are frequently observed. Under the influence of Turkmen patterns such as *gözüm, sözüm, kolum*, Kazakh written texts contain forms like *közüm, sözüm, qolum.*

c) Interference Related to Plural Markers

Kazakh employs six allomorphic variants of the plural marker *(-lar, -ler, -dar, -der, -tar, -ter*), whereas Turkmen uses only the forms *-lar* and *-ler*. This structural difference makes it difficult to select the appropriate plural marker in accordance with phonetic harmony.

Common erroneous forms include *sabaqlar, oryndyqdar, jetistikler, ağaşlar, and qalamlar*. These examples indicate a tendency toward automatic generalization based on a single morphological template and an insufficiently developed morphonological sensitivity.

c) Interference Related to Case Markers

Kazakh has seven grammatical cases, whereas Turkmen has six (*düşüm)* The instrumental case is absent from the Turkmen grammatical system; its functions are expressed through the postpositional construction *bilen*. Consequently, students frequently replace the Kazakh instrumental case markers *–men, –ben, –pen* with constructions containing *bilen*.

Thus, instead of the normative form *Arslan men Batyr oqidy*, students produce *Arslan bilen Batyr oqidy*. The use of *bilen* in place of instrumental case markers is systematically attested in students' written texts.

Moreover, students often fail to observe vowel harmony when selecting case markers and directly transfer analogical patterns from their native language, producing forms such as *injenernıñ, qyznyñ, mūğalımdıñ, mūğalımğa, kıtapda, ağaşdan*, among others.

Overall, these findings demonstrate insufficient morphonological adaptation of the Kazakh case system among Turkmen students and clearly reveal the key mechanisms underlying cross-linguistic morphological interference.

3)The Impact of Interference-Related Errors on the Cohesion and Coherence of Students' Written Discourse

The impact of interference on the genre of written summaries was measured by analyzing cohesion and coherence. The results indicate that the "cognate cost" (the effort to inhibit the L1) consumes the learner's attentional bandwidth, leading to fragmented academic discourse.

- The quantitative analysis revealed that subject-predicate agreement errors were among the most frequent interference-related deviations (n=247, 18.0% of total errors), directly disrupting chains of reference and cohesive links. In 73% of affected texts, these morphological deviations broke the referential chain across at least two consecutive sentences, reducing inter-sentential cohesion scores significantly.
- At the level of coherence, lexical interference and the use of "false friends" (lexical cognates with divergent meanings) resulted in semantic imprecision. The thematic progression of the summaries was frequently interrupted as students resorted to Turkmen-based idiomatic structures, thereby reducing the texts' "comprehensibility" and "task requirement" scores.

The findings of the analysis indicate that phonetic, graphemic, morphological, and lexical interference-related errors observed in the written discourse of Turkic-speaking students, including Turkmen learners, are not confined to isolated linguistic levels but have a direct and cumulative impact on the overall structural and semantic integrity of the text. These errors lead to a weakening of cohesion and coherence - core indicators of written discourse quality- and, consequently, diminish students' communicative competence in academic writing.

Interference was found to manifest most prominently at the level of cohesion. Morphological interference, particularly the incorrect or inappropriate use of personal, possessive, plural, and case affixes, undermines both intra-sentential and inter-sentential grammatical relations. Disruptions in agreement between subject and predicate in terms of person and number, as well as the incorrect selection of case forms, hinder the formal linking of sentences. As a result, cohesive chains within the text are formed inconsistently, making it difficult for readers to follow the progression of ideas.

Graphemic interference represents another significant factor contributing to the erosion of cohesion. The mixed use of Cyrillic and Latin scripts disrupts the text's visual integrity and impedes automatic word recognition. Consequently, readers must expend additional cognitive effort during text processing, which negatively affects overall comprehension. The lack of orthographic consistency thus emerges as an additional constraint on the continuity of cohesive relations.

Interference-related errors also have a pronounced effect on textual coherence. Morphological and lexical deviations impede the logical development of ideas by weakening the functioning of discourse markers that encode causal, temporal, and sequential relations. As a result, thematic progression is often only partially realized, and the flow of argumentation becomes fragmented and discontinuous. Insufficient logical connectivity between paragraphs further disrupts the text's macrostructural organization, hindering the coherent development of the central argument.

Phonetic and graphemic interference affect coherence indirectly. Misspellings and distorted phonological representations slow down text comprehension and obstruct access to core content. In academic texts in particular, such deviations reduce the perceived credibility and semantic transparency of the discourse. In many student texts, the logical relationship among the introduction, main body, and conclusion is weak, and genre conventions are not consistently observed. This tendency can be attributed to the direct transfer of text-construction patterns from the native language into Kazakh writing, as well as to an insufficient command of second-language discourse norms.

The qualitative analysis demonstrates that:

- Morphological errors weaken inter-sentential grammatical relations;
- Graphemic errors reduce the level of academic acceptability of written texts.

Overall, interference-related errors in students' written discourse result in systematic disruptions of cohesive devices, weaken the logical and semantic unity of texts, and prevent full compliance with academic and genre-specific conventions. These findings support the view that interference should be understood not merely as a linguistic error, but as a complex cognitive–discursive phenomenon that directly affects the quality of written communication. Accordingly, the systematic implementation of translanguaging and contrastive pedagogical approaches, aimed at the conscious management of interference, along with targeted tasks designed to foster written discourse development, is of particular importance in teaching Kazakh to Turkic-speaking learners.

The preceding section presented the corpus-based findings on interference-related error patterns and their cumulative impact on textual cohesion and coherence. The following discussion interprets these findings in relation to the theoretical framework and prior empirical literature, articulates the study's theoretical novelty and practical implications, and acknowledges its limitations. The findings of the present study indicate that interference in the written discourse of Turkic-speaking students, including Turkmen learners, is systematic and multilevel.

The analyzed data clearly demonstrate the stable and recurrent character of phonetic interference. In producing written texts in Kazakh, Turkmen students first commit orthographic errors resulting from mismatches between graphemes and phonemes; second, substitute phonetically similar sounds due to differences in articulatory bases; and third, are influenced by divergent alphabetic systems, which directly affect the formation of orthographic competence.

These patterns confirm that phonetic interference is not incidental but rooted in systematic cross-linguistic contrasts.

Table 2.

Model of the Impact of Interference on the Cohesion and Coherence of Students' Written Discourse

Level / Process	*Description*
Native language (L1) system	A fully established system of phonetic, morphological, lexical, graphemic, and cognitive patterns acquired during primary language socialization
Cross-linguistic transfer	The automatic activation and transfer of structural and cognitive patterns from the native language to the second language (L2) during language acquisition
Types of interference	Phonetic, graphemic, morphological, and lexical interference
Linguistic-level deviations	Distortions in sound–grapheme correspondences; incorrect affixation; misuse of false lexical equivalents; mixing of Cyrillic and Latin scripts
Disruption of cohesion	Breakdown of subject–predicate agreement; weakening of case and possessive relations; excessive and functionally unjustified lexical repetition; lack of orthographic consistency
Weakening of coherence	Disruption of thematic progression; attenuation of causal, temporal, and logical relations; absence of coherent paragraph transitions; failure to maintain the textual macrostructure (introduction–body–conclusion)
Decline in written discourse quality	Deviation from academic style conventions; reduced textual clarity and readability; incomplete realization of communicative and academic objectives

These findings underscore the importance of systematic and contrastive instruction in sound correspondences and graphemic distinctions when teaching Kazakh as a cognate language. To mitigate phonetic interference, the instructional process should incorporate targeted activities, including audiovisual dictation, tasks involving minimal pair comparisons, and articulatory training exercises designed to raise phonological awareness.

Interference-related errors at the morphological level are among the most frequent sources of difficulty for Turkic-speaking learners acquiring Kazakh. Their emergence can be attributed to several interrelated factors. First, reliance on superficial similarities between morphemic systems leads to false analogy. Second, the syncretic nature of plural and possessive forms in Turkmen complicates the differentiation of corresponding grammatical categories in Kazakh. Third, the absence of the instrumental case in Turkmen, combined with the presence of the functional construction *bilen*, results in the automatic substitution of Kazakh instrumental affixes *(–men, –ben, –pen*). Finally, incomplete mastery of vowel harmony rules contributes to the stabilization of morphonological errors.

To reduce the impact of these interference phenomena, the study proposes several pedagogically effective strategies, including the systematic use of contrastive grammatical exercises, the analysis of a corpus of learners' typical errors, and the integration of translanguaging-based and contrastive explanatory approaches into classroom instruction.

These instructional approaches contribute to the development of learners' metalinguistic awareness, promote conscious lexical selection, and ultimately reduce the incidence of lexical interference.

The phonetic, morphological, and graphemic interference-related errors identified in this study are not random, but rather stable and predictable. This regularity reflects both the typological proximity of Turkmen and Kazakh and the differences in their phonological, morphonological, and orthographic systems. At the phonetic level, substitutions resulting from articulatory differences are particularly frequent, whereas at the morphological level, errors based on false analogy and analogical overgeneralization predominate. These patterns constitute clear manifestations of negative cross-linguistic transfer.

Importantly, the effects of interference extend beyond the level of linguistic form to the structural integrity of written texts. Grammatical errors disrupt the appropriate use of cohesive devices, including reference, ellipsis, conjunction, and lexical repetition, thereby weakening internal textual links and reducing readability and logical continuity.

Interference is also clearly manifested at the level of coherence. In the written texts produced by Turkmen students, thematic progression is often weak, causal relations are insufficiently articulated, and genre conventions are inconsistently observed. This suggests that, during text production, learners rely heavily on semantic and cognitive models derived from their native language. Consequently, the discourse norms of Kazakh are not fully activated in processes of information selection, generalization, and reorganization, leading to disruptions in the overall logical structure of the text.

The results of the present study are consistent with and extend findings reported in prior research. Johanson (2002) documented the pervasive role of structural divergence in Turkic language contacts. The present corpus confirms that phonetic and morphological mismatches between Kazakh and Turkmen generate stable, predictable error patterns rather than random performance slips, aligning with Odlin's (1989) foundational claim that the degree of L1–L2 structural difference reliably predicts the incidence and type of transfer errors. At the graphemic level, the Cyrillic–Latin script mixing identified here extends Tastanbek's (2019) observation that digital communication environments reinforce L1 orthographic habits in Kazakhstani learners. Regarding discourse quality, the finding that morphological interference disrupts cohesive chains corroborates Halliday and Hasan's (1976) framework, which treats grammatical cohesion – including reference, conjunction, and ellipsis – as the primary scaffold of textual integrity. Where prior studies had reported isolated phonetic or morphological errors (Johanson, 2002; Odlin, 1989), the present study demonstrates that these deviations accumulate to erode the macrostructural coherence of students' written summaries, a discourse-level link not previously

established for Turkmen–Kazakh learner language. The finding that insufficiently developed L1 academic writing skills directly impairs L2 coherent text production reinforces Cummins' (1981) Common Underlying Proficiency proposition.

The theoretical novelty of this study lies in its reconceptualization of interference not merely as a set of isolated linguistic errors but as a cognitive-discursive phenomenon that operates simultaneously across multiple levels and accumulates to undermine the quality of discourse-level texts. By mapping the pathway from L1-triggered error patterns to breakdowns in cohesion and coherence, the study produces the first level-based typological classification of Turkmen–Kazakh interference in written academic summaries, which did not previously exist in the literature. The principal practical implication is that diagnostic and instructional tools for Kazakh as a cognate language must address not only form-level accuracy but also the discourse-structural consequences of systematic interference, particularly at the A1 level where transfer is most salient. A key limitation of the study is that it focuses exclusively on written discourse among A1-level learners at a single institution; future research should employ longitudinal designs, include oral production data, and extend sampling across proficiency levels and institutions to assess the generalizability of these findings.

From a practical standpoint, the findings offer important guidance for developing instructional materials for Turkic-speaking learners. Taking into account the systematic typology of interference-related errors, the study substantiates the need for contrastive grammatical exercises, tasks based on phonetic and graphemic comparison, and writing strategies explicitly aimed at developing cohesion and coherence. The adoption of a translanguaging approach enables learners' native language to function not as an obstacle but as a pedagogical resource, enhances metalinguistic reflection, and improves the overall quality of written production.

The present study demonstrates that the influence of the native language in the written discourse of Turkic-speaking students, including Turkmen learners acquiring Kazakh as a foreign language, is systematic and multi-level. The findings confirm that interference manifests at phonetic, morphological, lexical, graphemic, and discursive levels and constitutes one of the key factors undermining textual cohesion and coherence. In particular, morphological interference was found to be highly persistent, resulting in disruptions of inter-sentential grammatical relations due to the incorrect use of case, possessive, plural, and personal affixes. At the lexical level, cognate roots and false cognates are frequently misinterpreted due to cognitive overgeneralization, leading to semantic imprecision.

The study substantiates the necessity of conceptualizing interference not merely as a set of formal linguistic errors, but as a cognitive-discursive phenomenon that directly affects textual integrity. Disruptions in cohesive devices such as reference, conjunctions, and ellipsis, together with weakened thematic progression, lead to violations of logical consistency and genre conformity in students' written texts.

CONCLUSION

The scholarly contribution of this research lies in developing a level-based, typological classification of interference-related errors characteristic of Turkmen students' written Kazakh, as well as in empirically validating the effectiveness of translanguaging and contrastive instructional approaches using authentic linguistic data. Treating the native language not as an obstacle but as a cognitive resource that promotes metalinguistic reflection provides a principled foundation for improving the methodology of teaching Kazakh as a cognate language.

At the same time, the study has certain limitations. It focused exclusively on written discourse and did not include data on oral production. Future research should therefore aim to expand the multilevel corpus of interference-related errors, conduct longitudinal investigations, and employ cognitive and neurolinguistic methods to uncover the underlying mechanisms of cross-linguistic transfer. Additionally, the development of automated tools for error diagnosis and correction using digital technologies may provide an effective means of enhancing the written competence of Turkic-speaking students in Kazakh.

Overall, the findings substantiate the need to implement a translanguaging-based and contrastive methodological model oriented toward the conscious management of interference in teaching Kazakh as a cognate language. The study thus provides a solid theoretical and practical foundation for further research and pedagogical innovation in this domain.

Information about funding

This research was funded by a grant from the Ministry of Education and Science of the Republic of Kazakhstan (Project No. AP23487324, Teaching the Kazakh Language to Turkic-Speaking Students in a Translingual Aspect Based on Digital Technologies, 2024–2026).

REFERENCES

Amaniyazova, A. (2020). *Kazakhstani teachers' beliefs on translanguaging: evidence from a trilingual context*. Nazarbayev University, Kazakhstan. http://nur.nu.edu.kz/handle/123456789/4853

Brezina, V. (2018). *Statistics in corpus linguistics: A practical guide*. Cambridge University Press. https://doi.org/10.1017/9781316410899

Broersma, M., & De Bot, K. (2006). Triggered codeswitching: A first inventory of shared and non-shared clues. *International Journal of Bilingualism*, *10*(3), 337–361.

Canagarajah, S. (2011). Translanguaging in the classroom: Emerging issues for research and pedagogy. *Applied Linguistics Review*, *2*(1), 1–28. https://doi.org/10.1515/9783110239331.1

Clyne, M. (2003). *Dynamics of language contact: English and immigrant languages*. Cambridge University Press. https://doi.org/10.1017/CBO9780511606526

Corder, S. P. (1967). The significance of learner's errors. *International Review of Applied Linguistics in Language Teaching*, *5*(1-4), 161–170. http://dx.doi.org/10.1515/iral.1967.5.1-4.161

Corder, S. P. (1975). Error analysis, interlanguage and second language acquisition. *Language Teaching & Linguistics: Abstracts*, *8*(4), 201–218. https://doi.org/10.1017/S0261444800002822

Creswell, J. W., & Creswell, J. D. (2017). *Research design: Qualitative, quantitative, and mixed methods approaches* (5th ed.). SAGE Publications. https://doi.org/10.1080/15424065.2022.2046231

Cummins, J. (1981). The role of primary language development in promoting educational success for language minority students. In California State Department of Education (Ed.), *Schooling and language minority students: A theoretical framework* (pp. 3–49). Evaluation, Dissemination and Assessment Center. https://doi.org/10.13140/2.1.1334.9449

Dijkstra, A., Grainger, J., & van Heuven, W. J. B. (1999). Recognition of cognates and Interlingual Homographs: The Neglected Role of Phonology. *Journal of Memory and Language*, *41*(4), 496–518. http://doi.org/10.1006/jmla.1999.2654

García, O., & Li, W. (2014). *Translanguaging: Language, bilingualism and education*. Palgrave Macmillan. https://doi.org/10.1057/9781137385765

García, O., & Kleifgen, J. A. (2020). Do languages exist? Translanguaging and its consequences for bilingual education. In V. Vaish (Ed.), *Translanguaging in multilingual English classrooms* (pp. 13–30). Springer Nature Singapore. https://doi.org/10.1007/978-981-15-0912-4_2

Goodman, B., & Tastanbek, S. (2020). Making the state language a "prestige" language? A policy analysis of Kazakh as a medium of instruction in Kazakhstan. *Language Policy*, *20*, 485–509.

Granger, S. (2003). Error-tagged learner corpora and second language acquisition. *International Journal of Corpus Linguistics*, *8*(1), 123–146.

Halliday, M. A. K., & Hasan, R. (1976). *Cohesion in English*. Longman.

Imankulova, M., Nurgaziyev, T. & Moldagali, B. (2025). Effectiveness of Language Education in Kazakhstan: Integration of the Kazakh Language with Technology, Pedagogy, and Content (KFL-TPACK). *Journal of Social Studies Education Research, 16*(3), 332-348. https://jsser.org/index.php/jsser/article/view/6857/751

Ismagulova, A., Rakisheva, G., Chernigovtseva, O., Sabitova, D., & Gerfanova, E. (2024). Technology for Creating a Language Learning Environment for Teaching English. *Journal of Curriculum Studies Research, 6*(2), 234-258. https://doi.org/10.46303/jcsr.2024.20

Jarvis, S., & Pavlenko, A. (2008). *Cross-linguistic influence in language and cognition*. Routledge. ISBN 9780415879811

Johanson, L. (2002). *Structural factors in Turkic language contacts*. London, UK; New York, NY: Routledge. ISBN 9781138983076

Karabassova, L., & San Isidro-Smith, A. (2020). Implementing content and language integrated learning (CLIL) in Kazakhstan: Teachers' views. *International Journal of Bilingual Education and Bilingualism*, *23*(8), 1017–1031.

Karabulut, A., & Kesli Dollar, Y. (2022). The use of translanguaging pedagogy in writing classes of Turkish EFL learners. *Participatory Educational Research*, *9*(6), 41-65. https://doi.org/10.17275/per.22.128.9.6

Karlinsky, A. E. (1996). *Vzaimodeistvie yazykov: teoreticheskie i prikladnye aspekty*. Gylym.

Kaiypova, F., Lee, J. H., Jeon, J., & Lee, S. (2025). *Translanguaging and motivational development in CLIL: A case from Kazakhstan's trilingual context*. *International Journal of Applied Linguistics*. https://doi.org/10.1111/ijal.70031

Kuzembayeva, G., Maigeldiyeva, Z., Kussanova, B. & Maydangalieva, Z. (2024). Heritage Language Experiences and Proficiency Among Ethnic Kazakhs Abroad: Exploring Challenges, Language Skill Requirements, and Acquisition Barriers, *Journal of Social Studies Education Research, 15*(3), 348-373. https://jsser.org/index.php/jsser/article/view/5655/690

Levelt, W. J. M. (1989). Speaking: From intention to articulation. MIT Press. https://doi.org/10.7551/mitpress/6393.001.0001

McEnery, T., & Brookes, G. (2024). *The Routledge handbook of corpus linguistics* (2nd ed.). Routledge.

Odlin, T. (1989). Language transfer: Cross-linguistic influence in language learning. Cambridge University Press. http://dx.doi.org/10.1017/CBO9781139524537

Ringbom, H. (2006). *The importance of cross-linguistic similarity in foreign language learning*. Multilingual Matters.

Selinker, L. (1972). Interlanguage. *International Review of Applied Linguistics in Language Teaching*, *10*(3), 209–231. https://doi.org/10.1515/iral.1972.10.1-4.209

Stefanowitsch, A. (2020). Corpus linguistics: A guide to the methodology. Language Science Press. https://doi.org/10.5281/zenodo.3735822

Tastanbek, S. (2019). Translanguaging as a pedagogical resource in the university Kazakh language classroom. *Vestnik KarGU*, *1*(93), 201–208.

van Dijk, T. A. (1980). *Macrostructures: An interdisciplinary study of global structures in discourse, interaction, and cognition*. Lawrence Erlbaum Associates.

Wang, J., Hanafi Zaid, Y., Ibrahim, N. A., & Haladin, N. B. (2024). Teachers' and students' beliefs about translanguaging in higher education: A systematic review. *Journal of Nusantara Studies (JONUS)*, *9*(2), 402–419. https://doi.org/10.24200/jonus.vol9iss2pp402-419

Weinreich, U. (1953). *Languages in contact: Findings and problems*. Linguistic Circle of New York.

Williams, C. (1994). Arfarniadau dwyieithog yn addysg bellach a phrifysgol. Unpublished doctoral dissertation. University of Wales, Bangor.