ISSN 1799-2591 (print); ISSN 2053-0692 (online)
Theory and Practice in Language Studies, Vol. 16, No. 11, pp. 3937-3949, November 2026
DOI: https://doi.org/10.17507/tpls.1611.26

Grapheme–Morphonological Difficulties of English-Speaking Learners at the Initial Stage of Learning Kazakh: Toward Justifying an Independent Pre-A1 Level

Anar Fazylzhan
A. Baitursynuly Institute of Linguistics, Almaty, Kazakhstan

Didar Sadyk*
Al-Farabi Kazakh National University, Almaty, Kazakhstan

Madina Abayeva
A. Baitursynuly Institute of Linguistics, Almaty, Kazakhstan

Yelimay Shynatay
A. Baitursynuly Institute of Linguistics, Almaty, Kazakhstan

***Abstract*—This study is relevant because the earliest stage of Kazakh as a foreign language remains insufficiently described for English-speaking learners, whose initial acquisition involves not only elementary vocabulary and grammar, but also entry into a new script, unfamiliar sound categories, vowel harmony, and affixal allomorphy. This article aims to justify the need for an independent pre-threshold Pre-A1 stage in Kazakh instruction and to identify the grapheme-phonological and morphonological difficulties that hinder learners' transition to A1. The study is based on an eight-week intensive A1 Kazakh course at the University of Oxford, involving ten native speakers of English across three consecutive groups, and combines structured pedagogical observation, written assignments, audio-supported phonetic exercises, contrastive analysis of the 42 graphemes of the Kazakh Cyrillic alphabet, and qualitative interpretation of recurrent learner errors. The methodological contribution of the study lies in classifying Kazakh Cyrillic graphemes into familiar, deceptive, and unfamiliar groups for Latin-script users and operationalizing this classification through four coefficients: Visual Familiarity Coefficient, Direct Transfer Coefficient, Misfire Rate, and Interference Risk Coefficient. The results show that visual accessibility does not guarantee successful transfer: only 10 of 42 graphemes allow relatively safe direct transfer, while deceptive graphemes create a high risk of orthographically mediated phonological interference. The classroom evidence further shows that substitutions such as /q/ → /k/, /ʁ/ → /g/, and /ө/ → /y/ affect not only pronunciation, but also the recognition of vowel harmony and suffix selection. The findings substantiate a separate Pre-A1 module organized around script orientation, contrastive sound-letter mapping, perceptual training, vowel-harmony recognition, controlled affixal variation, and embedded correction.**

***Index Terms*—Kazakh as a foreign language, Pre-A1, grapheme-phonological interference, Kazakh Cyrillic, morphonology**

I. Introduction

Kazakh as a foreign language is gaining international relevance, but its earliest stage of acquisition remains less described than comparable stages in more widely taught languages. For English-speaking learners, the beginning of Kazakh study requires simultaneous entry into a new script, new sound categories, agglutinative word formation, vowel harmony, and affixal alternation. These domains interact: unstable decoding weakens perception, repetition, word classification, and inflection, while unstable sound categories complicate later grammatical choice. The pre-threshold stage therefore requires independent pedagogical attention rather than treatment as a shortened version of A1.

This problem is supported by intersecting research traditions. The CEFR Companion Volume recognizes Pre-A1 as a level of highly restricted but meaningful language activity before full A1 performance (Council of Europe, 2020). In Kazakh language pedagogy, recent studies show that existing resources and explanations remain insufficient for learners who enter Kazakh through English rather than Russian or another Turkic language (Zhumashova et al., 2023; Rsaliyeva & Chukayeva, 2025). Broader work on orthographic input demonstrates that writing systems shape phonological awareness, phoneme perception, and lexical encoding even when instruction seems mainly oral (Hayes-Harb & Barrios, 2021; Shepperd, 2024). L2 speech models and cross-linguistic similarity research further explain why English-speaking

© 2026 ACADEMY PUBLICATION

learners rely on familiar visual and acoustic categories when processing unfamiliar forms (Best & Tyler, 2007; Ringbom & Jarvis, 2011; Flege & Bohn, 2021).

Previous studies provide several foundations for the present work. Research on second language writing systems shows that orthography can affect phonological representations, lexical learning, and perception-production relations (Cook & Bassetti, 2005; Bassetti, 2006; Hayes-Harb et al., 2010). Studies of Russian Cyrillic among English-speaking learners demonstrate that the most persistent errors often arise from graphically familiar but phonologically misleading letters rather than from entirely unfamiliar symbols (Brennan & Kiskin, 2020; Showalter, 2018, 2020). L2 speech models explain non-native sound assimilation and the need for perceptual training before stable production, while Kazakh and Turkic phonological studies identify vowel harmony as a central structural principle rather than a minor pronunciation detail (Flege & Bohn, 2021; Bazarbayeva, 2012; McCollum, 2018).

A specific gap remains at the intersection of these findings. Orthographic transfer, L2 phoneme perception, pronunciation instruction, Cyrillic learning, Kazakh vowel harmony, and agglutinative morphology are usually examined separately. Yet English-speaking beginners encounter them together before they can operate at A1. Russian Cyrillic findings cannot be transferred directly to Kazakh because Kazakh Cyrillic includes graphemes such as ә, ө, ү, ұ, қ, ғ, ң, and і, which represent language-specific contrasts and participate in vowel-harmony-based morphology. The underdescribed issue is therefore not simply alphabet learning or pronunciation difficulty, but their interaction with early grammatical choice.

The novelty of this study lies in conceptualizing early difficulties as a unified pre-threshold risk zone. The article links script instruction and morphonology as sequential stages of initial Kazakh acquisition, classifies the 42 Kazakh Cyrillic graphemes by visual and phonemic similarity for Latin-script users, and operationalizes the classification through four coefficients: Visual Familiarity Coefficient, Direct Transfer Coefficient, Misfire Rate, and Interference Risk Coefficient. It also refines orthographically mediated phonological interference (OOPHI) as a mechanism in which a familiar-looking grapheme activates an L1-based sound routine before the Kazakh rule is applied. In this way, Pre-A1 is justified through observable learner behavior and measurable alphabetic risk.

The aim of the article is to justify an independent Pre-A1 stage in teaching Kazakh to English-speaking learners by identifying grapheme-phonological and morphonological difficulties that hinder the transition to A1. The study clarifies the role of orthographic input and L2 speech perception, analyzes risk zones in Kazakh Cyrillic, connects grapheme-phonological errors with vowel harmony and affixal variation, and formulates principles for a Pre-A1 module that reduces decoding load and separates perception failure from rule failure.

II. Literature Review

The literature review follows four analytical domains: Kazakh as a foreign language, orthographic input and cross-script transfer, L2 speech perception, and Kazakh morphonology. This organization reflects the learner's path from written form to sound-letter mapping, phoneme categorization, vowel-harmony recognition, and suffix selection.

Key terms are used in a narrow diagnostic sense. Pre-A1 denotes restricted language activity requiring support for decoding, perception, and elementary form-function mapping (Council of Europe, 2020). Grapheme-phonological interference refers to errors caused by mismatch between letter shape and target sound value. Orthographically mediated phonological interference is the activation of an incorrect L1-based sound by a visually familiar L2 grapheme. Morphonological difficulty arises when perception affects morphological choice, as in vowel-harmony-based suffix selection.

Table 1 summarizes the literature domains used to justify an independent Pre-A1 stage for Kazakh.

Table 1
Analytical Domains of the Literature on Pre-A1 Kazakh Instruction for English-Speaking Learners

Domain	*Core issue*	*Representative sources*	*Function in the present study*
Kazakh as a foreign language	Underdeveloped description of the English-speaking beginner's pathway	Ayapova & Thomas, 1992; Batayeva, 2012; Muhamedowa, 2016; Zhumashova et al., 2023	Defines the local and pedagogical need for a Pre-A1 module
Orthographic input and cross-script transfer	Influence of written form on phonological learning and lexical encoding	Cook & Bassetti, 2005; Hayes-Harb & Barrios, 2021; Shepperd, 2024; Showalter, 2020	Explains why visual similarity can be a source of error rather than support
L2 speech perception and pronunciation instruction	Formation of new phoneme categories and perception-production links	Best & Tyler, 2007; Flege & Bohn, 2021; Lee et al., 2015; Saito & Plonsky, 2019	Provides theoretical and instructional grounding for early phonetic diagnosis
Kazakh vowel harmony and morphonology	Interaction between phonemic perception, harmonic classes, and affixal allomorphy	Bazarbayeva, 2012; McCollum, 2018; Zhunisbek, 2018; Rsaliyeva & Chukayeva, 2025	Connects the alphabetic barrier with grammatical development

A. Kazakh as a Foreign Language and the Pre-A1 Problem

© 2026 ACADEMY PUBLICATION

Teaching Kazakh to non-native learners has a long tradition, but it developed unevenly across learner groups. Early and later instructional traditions mainly served Russian-speaking audiences, whose literacy background and grammatical expectations differ substantially from those of English-speaking learners. For the latter group, Kazakh is encountered not only as a new lexical and grammatical system but also as a new alphabetic, phonological, and typological system. This distinction is central because transfer from Russian-based materials cannot automatically solve problems created by English literacy and English phonology.

English-language resources have expanded through textbooks, grammars, and reference materials (Ayapova & Thomas, 1992; Batayeva, 2012; Krippes, 1996; Muhamedowa, 2016). However, these resources do not fully define the micro-stage before A1, when learners are still learning to decode script, map letters to sounds, identify harmonic classes, and understand suffix alternation. Recent studies confirm the need for adapted materials and corpus-based approaches for anglophone learners (Zhumashova et al., 2023; Rsaliyeva & Chukayeva, 2025). The English-speaking beginner therefore cannot be treated as a simple variant of the Russian-speaking beginner.

Kazakh learning also occurs within a multilingual educational ecology shaped by revitalization, trilingual education, and language policy challenges (Kucherbayeva & Smagulova, 2023; Sarmurzin et al., 2024; Yelubayeva et al., 2025). This broader context matters because foreign-language instruction in Kazakh is tied to the international visibility of the language and to the development of specialized materials. Recent studies in Theory and Practice in Language Studies further show that Kazakh material requires models combining structural, pedagogical, corpus-based, and intercultural interpretation (Bakytzhanova et al., 2025; Ormanova et al., 2025). Related work on Turkic terminology and university-level teaching also emphasizes the importance of systematic linguistic organization and technology-supported instruction in Turkic-language education (Zhanzhigitov et al., 2022).

The Pre-A1 problem is therefore both linguistic and instructional. CEFR descriptors acknowledge the stage before A1, but they do not specify the language-specific burden of Kazakh (Council of Europe, 2020). For English-speaking learners, this burden includes Cyrillic recognition, new phonemes, vowel-based suffix choice, and agglutinative compression. The form dostarymyzdykindemiz (a complex possessive-locative-predicative form built on dostar “friends”) illustrates why segmentation and morphonology must be prepared before full A1 progression.

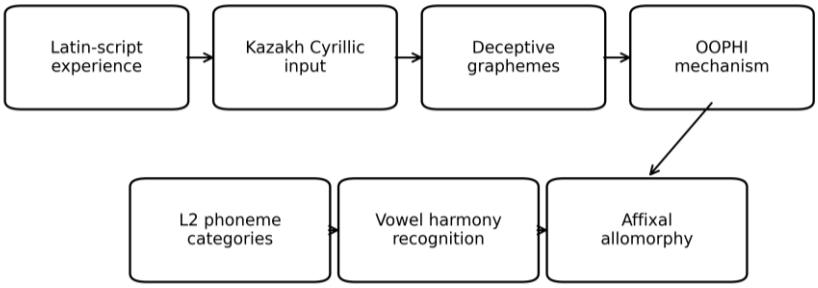


Figure 1. Conceptual Architecture of Pre-Threshold Risk in Initial Kazakh Acquisition

Figure 1 shows that learners do not move harmlessly from script to grammar. Latin-script experience shapes the perception of Kazakh Cyrillic; deceptive graphemes may trigger OOPHI; and this may affect phoneme categorization, vowel harmony, and affixal allomorphy. A Pre-A1 module must therefore address the alphabetic barrier as a morphonological risk.

B. Orthographic Input, Cross-Script Transfer, and Graphemic Interference

Research on second language writing systems rejects the view of orthography as a neutral reflection of speech. Written input can shape phonological representations, speech production, and new word learning (Cook & Bassetti, 2005; Bassetti, 2006; Erdener & Burnham, 2005; Hayes-Harb et al., 2010). In Kazakh, beginners often meet written and spoken forms together, so script can both support and mislead acquisition.

Recent reviews confirm that orthography affects phonological awareness, phoneme perception, syllable structure, and word-form learning in L2 acquisition (Hayes-Harb & Barrios, 2021; Shepperd, 2024). Learners therefore process Kazakh Cyrillic through prior literacy routines rather than as a neutral symbol set. Because Kazakh Cyrillic includes letters for phonemes absent from English, visual similarity must be treated as a measurable risk. The transferability of Cyrillic-Latin handwriting skills also indicates that script learning has sensorimotor and cognitive dimensions (Asselborn et al., 2021). For the present article, the Jaccard similarity principle supports formal comparison of overlap between grapheme sets (Jaccard, 1901).

Russian Cyrillic studies are useful but not directly transferable. Showalter (2018, 2020) and Brennan and Kiskin (2020) show that familiar-looking but misleading letters cause persistent errors in Russian. Kazakh expands this problem because қ, ғ, ң, ә, ө, ұ, ү, and і represent Kazakh-specific phonological contrasts and participate in vowel-harmony-based morphology.

The distinction between familiar, deceptive, and unfamiliar graphemes is therefore central. Familiar graphemes can support positive transfer; deceptive graphemes create a misfire effect by activating an incorrect sound value; unfamiliar

© 2026 ACADEMY PUBLICATION

graphemes require memorization but do not necessarily trigger misleading L1 pronunciation. Misfire Rate captures this risk among familiar-looking symbols.

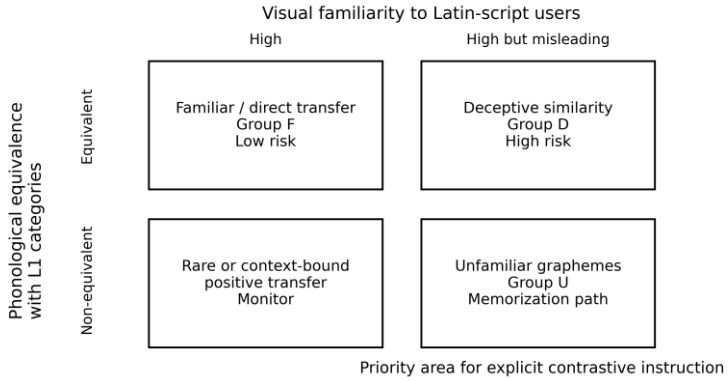


Figure 2. Cross-Script Interference Matrix for English-Speaking Learners of Kazakh Cyrillic

Figure 2 converts orthographic-transfer theory into a diagnostic matrix. The high-risk zone is not low familiarity alone, but visual familiarity without phonological equivalence. This explains why Pre-A1 instruction should include explicit contrastive pairs: how a letter looks and how it is pronounced.

C. L2 Speech Models, Pronunciation Instruction, and Diagnostic Sequencing

L2 speech models explain why some Kazakh sounds are difficult for English-speaking learners. SLM and SLM-r relate acquisition to the distance between target-language categories and L1 categories, while PAM-L2 specifies perceptual assimilation to the closest L1 category (Best & Tyler, 2007; Flege, 1995; Flege & Bohn, 2021). Kazakh /q/, /ʁ/, /ɯ/, /ø/, and /ʏ/ are therefore not peripheral pronunciation details but contrastive categories affecting lexical and morphonological development.

The present data extend these models to script-switching conditions. A deceptive Cyrillic letter can activate a Latin-script or English-based phonological routine before the target sound is processed. Thus, OOPHI functions as a written-input route into perceptual assimilation: the error may begin not only with hearing, but with the way sound is visually cued.

Pronunciation-instruction research supports early intervention. Meta-analyses show that instruction can improve L2 pronunciation, especially when perception, intelligibility, and communicative function are prioritized rather than isolated imitation (Saito, 2012; Lee et al., 2015; Saito & Plonsky, 2019; Thomson & Derwing, 2015; Derwing & Munro, 2015). Studies of orthography and non-native vowel contrasts also show that written forms can influence distinctions later used in morphology (Simon et al., 2010). Applied to Kazakh, this means that Pre-A1 phonetic work should connect sounds to graphemes and to grammatical consequences such as suffix selection.

This literature justifies the exploratory case-study design. Small-sample classroom data are analytically valuable when the aim is to identify mechanisms in context rather than estimate population frequencies. The Oxford course is treated as a high-information case in which typological distance, a new script, and intensive instruction make early difficulty visible (Yin, 2018).

The pedagogical consequence is diagnostic sequencing. If a learner selects the wrong suffix, the cause may be rule failure, incorrect perception of the root vowel, unstable grapheme recognition, or articulatory difficulty. Corrective feedback must identify the level of error before adding grammatical explanation.

D. Vowel Harmony, Morphonology, and the Design of a Pre-A1 Module

Kazakh is an agglutinative Turkic language in which grammatical meanings are expressed through suffixes attached to a stem. Unlike English, it can compress plurality, possession, case, relational meaning, and predication into one word form. This structural difference has a direct effect on early pedagogy. A learner cannot postpone morphonology until after vocabulary and grammar have been acquired, because even elementary forms require recognition of suffix variants. The example dostarymyzdykindemiz (a complex possessive-locative-predicative form built on dostar “friends”) illustrates why segmentation and harmonic recognition must be introduced carefully at the beginning.

Vowel harmony links sound perception with suffix selection. McCollum (2018) shows that Kazakh harmony is a structured phonological system, and comparative Turkic work confirms the importance of positional and phonetic factors (McCollum, 2020). Kazakh descriptions also stress front/back contrasts and language-specific vowels (Bazarbayeva, 2012; Muhamedowa, 2016; Zhunisbek, 2018). Because English lacks productive vowel harmony, the rule cannot be acquired through direct L1 analogy. It must be made visible, audible, and operational through examples and controlled practice.

© 2026 ACADEMY PUBLICATION

Vocabulary research reinforces this point: knowing a word includes form, meaning, use, and grammatical behavior (Nation, 2001; Webb & Nation, 2017). In Kazakh, form includes harmonic class and suffix behavior. A learner who knows only the stem has not yet learned the word as a usable unit.

A Pre-A1 module should therefore move from script orientation and warning about deceptive graphemes to contrastive sound-letter mapping, perceptual training, vowel-harmony recognition, transparent affix families, and embedded correction. This sequence reduces decoding load before grammatical complexity increases.

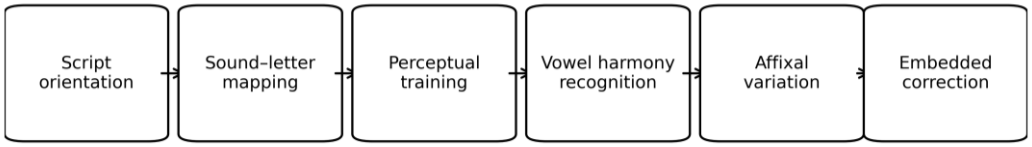


Figure 3. Literature-Based Sequencing of a Kazakh Pre-A1 Module

Figure 3 synthesizes this sequence. It begins with script orientation because written decoding is the first barrier for Latin-script users, but it does not stop there. The sequence moves to sound-letter mapping, perceptual training, vowel-harmony recognition, affixal variation, and embedded correction. This progression also explains the pedagogical usefulness of VFC, DTC, MFR, and IRC: they help determine which graphemes need brief familiarization, which require explicit warning, and which should be memorized without misleading analogy.

The reviewed literature converges on one conclusion: early Kazakh learning by English-speaking learners involves a compound difficulty. Kazakh pedagogy identifies the need for adapted materials; orthographic research explains why written input can support or distort phonological learning; L2 speech models explain category assimilation; pronunciation pedagogy supports early intervention; and Turkic phonology shows why vowel harmony and allomorphy must begin early. None of these traditions alone fully explains the transition from script to morphonology.

The present article addresses the remaining intersection: how Latin-script processing of Kazakh Cyrillic affects sound categories and morphonological decisions. By connecting classroom observation with grapheme classification, the study treats difficult letters, phonemes, and suffixes as one diagnostic chain. This chain moves from deceptive visual similarity to incorrect sound value, from unstable sound category to distorted vowel classification, and from distorted classification to allomorph-selection error. The model therefore does not merely list difficult elements of Kazakh; it explains why they become consequential at the same moment in the learner's trajectory. This is the point at which the literature review leads directly to the methodology: pre-threshold risk must be detected, measured, and transformed into course-design principles.

III. Methods

The study was designed as a qualitative-quantitative exploratory case study of pre-threshold difficulties in initial Kazakh acquisition by English-speaking learners. It combines structured pedagogical observation, written learner output, audio-supported phonetic exercises, and contrastive analysis of the 42 Kazakh Cyrillic graphemes.

The design links observable learner behavior with measurable properties of the input system. Classroom data identify recurring problems in decoding, pronunciation, and suffix selection, while grapheme-based coefficients operationalize alphabetic risk. The Oxford course is therefore used as a high-information case, not a representative population sample (Yin, 2018).

A. Materials and Participants

The empirical basis was an eight-week intensive A1 Kazakh course at the University of Oxford, four academic hours per week, across three consecutive groups. The ten participants were native speakers of English with no prior systematic Kazakh learning. They included students whose research projects concerned Kazakhstan, researchers and professionals in Central Asian studies, language enthusiasts without direct academic ties to the region, and heritage learners of Kazakh origin raised in the United Kingdom. This heterogeneity limits demographic generalization but strengthens the analytical value of the case because similar early difficulties appeared across different motivations and backgrounds.

The material included structured observation notes, written assignments, audio recordings of phonetic exercises, and the 42-grapheme Kazakh Cyrillic inventory. Observation recorded the target form in orthography and IPA transcription, learner production, linguistic context, and recurrence across the group. Written tasks traced grapheme recognition, segmentation, and suffix selection; audio recordings supported identification of recurring phonological substitutions; the grapheme inventory enabled classification by visual familiarity and phonemic correspondence for Latin-script users. Table 2 summarizes the data sources and their analytical functions.

© 2026 ACADEMY PUBLICATION

TABLE 2
EMPIRICAL MATERIAL AND ANALYTICAL FUNCTION OF THE STUDY

Data source	*Scope confirmed by the manuscript*	*Unit of analysis*	*Analytical function*
Structured pedagogical observation	Eight-week Oxford A1 course; three consecutive groups; ten English-speaking learners	Session-level observation record; target form; learner production; linguistic context	Identification of recurring pre-threshold difficulties in script recognition, sound-letter mapping, vowel harmony, and affixal variation
Written assignments	Learner-produced written forms collected during the course	Grapheme use, word form, suffix choice, and segmentation pattern	Detection of errors that appear in written decoding and elementary morphonological production
Audio recordings of phonetic exercises	Recordings of controlled pronunciation and repetition activities	Instructor IPA transcription of target and learner-produced forms	Support for identifying recurring phonological substitutions and perception-production mismatches
Kazakh Cyrillic grapheme inventory	42 graphemes classified into familiar, deceptive, and unfamiliar groups	Grapheme group and risk category	Calculation of VFC, DTC, MFR, and IRC as operational indicators of cross-script interference risk
Suffix and allomorph patterns	Vowel-harmony and consonant-conditioned examples from elementary Kazakh forms	Stem vowel class, final consonant type, and suffix variant	Diagnosis of the transition from phonological recognition to morphonological selection

The material demonstrates triangulation across pedagogical, written, audio-supported, and system-level evidence. Observation provided the instructional context, written assignments revealed how learners processed visible word forms, audio recordings supported interpretation of pronunciation-related errors, and the grapheme inventory translated qualitative difficulty into measurable indicators. This organization also responds to the need to connect the stated methodology with the actual analysis. Instrumental acoustic measurements were not conducted; therefore, phonological claims are limited to recurring classroom patterns documented through instructor IPA transcription and interpreted through L2 speech-learning models.

B. Methods and Instruments

The study integrates structured pedagogical observation, contrastive grapheme analysis, qualitative error analysis, and coefficient-based risk modeling. Observation documented difficulties with Cyrillic, sound-letter correspondences, Kazakh-specific phonemes, vowel harmony, and early affixal variation in instructional interaction.

Contrastive grapheme analysis compared the Latin alphabet familiar to English-speaking learners with Kazakh Cyrillic. The 42 graphemes were grouped as familiar, deceptive, or unfamiliar according to visual and phonemic correspondence. The Jaccard similarity principle made overlap expressible as a ratio rather than an impressionistic judgment and supported calculation of VFC, DTC, MFR, and IRC (Jaccard, 1901). This step was necessary because the main problem is not unfamiliarity alone, but the relation between visible similarity and sound equivalence.

Qualitative error analysis examined only recurring patterns across sessions and groups, not isolated slips. It focused on substitutions involving /q/ → /k/, /ʁ/ → /g/, and /ø/ → /y/, as well as difficulties with the back/front vowel opposition used in suffix selection. Examples were given in Latin transliteration with translations, such as qalada (“in the city”), ğalym (“scholar”), änşi (“singer”), balyqşy (“fisherman”), mektep-ke (“to the school”), and qala-ğa (“to the city”). This format preserved the Kazakh structure while making the examples accessible to international readers.

The interpretation of substitutions was guided by SLM, SLM-r, and PAM-L2, which explain the assimilation of non-native sounds to existing L1 categories (Best & Tyler, 2007; Flege & Bohn, 2021). Combined with orthographic-transfer research, these models support the concept of OOPHI as a script-triggered route into phoneme misclassification.

AI-assisted and digital tools were used only for auxiliary tasks: proofreading, academic wording, terminology consistency, bibliographic organization, transliteration checking, DOI/URL formatting, and table presentation. They were not used to fabricate data, invent learner productions, generate frequency counts, replace IPA-based analysis, or make unsupported interpretations. All classifications, calculations, examples, and conclusions remained subject to human scholarly control.

Table 3 links each method to its data source, analytical operation, and resulting output.

© 2026 ACADEMY PUBLICATION

TABLE 3
CORRESPONDENCE BETWEEN DATA, METHODS, AND RESULTS

Methodological component	*Data source*	*Analytical operation*	*Resulting output*
Structured pedagogical observation	Session notes, written tasks, and audio-supported phonetic exercises	Recording target forms, learner productions, contexts, and recurrent difficulties	Diagnostic profile of early grapheme-phonological and morphonological difficulties
Contrastive grapheme analysis	Latin alphabet and 42 Kazakh Cyrillic graphemes	Grouping graphemes into familiar, deceptive, and unfamiliar categories	Classification of interference risk zones for Latin-script users
Coefficient-based modeling	F, D, U groups of the Kazakh Cyrillic alphabet	Calculating VFC, DTC, MFR, and IRC	Quantified indicators of visual familiarity, safe transfer, misfire risk, and overall interference risk
Qualitative error analysis	Target forms and learner productions documented during the course	Separating graphemic, phonological, vowel-harmony, and affixal sources of error	Interpretation of recurring substitutions and suffix-selection difficulties
Pedagogical modeling	Results of diagnostic analysis and Pre-A1 learning constraints	Transforming risk indicators into instructional priorities	Principles for a separate Pre-A1 block before full A1 progression

The methodological design is cumulative: classroom evidence is formalized through grapheme grouping and coefficients, interpreted through L2 speech and orthographic-transfer theory, and converted into Pre-A1 course-design principles. This sequence prevents the recommendations from appearing as general teaching advice detached from the analyzed material and makes clear how each data type contributes to the results.

The design is exploratory but controlled. It is exploratory because the sample is small and context-specific; it is controlled because the claims are tied to documented classroom evidence, a fixed grapheme inventory, explicit formulas, and a defined interpretive procedure. Broader generalization requires larger samples and instrumental acoustic validation.

IV. RESULTS

The 42 Kazakh Cyrillic graphemes were classified according to their visual and phonemic similarity to Latin-script forms (Shepperd, 2024).

The results are presented in four steps: the alphabet-level interference profile, the calculated coefficients, recurrent classroom error patterns, and vowel harmony with affixal allomorphy as the morphonological continuation of the initial grapheme-phonological barrier. This order follows the learner's movement from written input to sound categorization and then to grammatical choice.

The evidence allows quantitative documentation at the alphabet level and qualitative documentation at the classroom level. Coefficients show system-level risk; learner errors show how this risk appears in instruction. Because no acoustic measurements were conducted, the results do not claim instrumental phonetic validation, but they do document recurring patterns through observation, written work, and instructor IPA transcription. This distinction is important for interpreting the results responsibly: the study measures alphabetic risk directly and treats phonological patterns as documented pedagogical evidence rather than as acoustic data.

A. Grapheme-Level Distribution of Interference Risk

The first result concerns the writing system. For Latin-script users, Kazakh Cyrillic is uneven: some letters support direct transfer, some are deceptive, and others require memorization.

TABLE 4
CLASSIFICATION OF KAZAKH CYRILLIC GRAPHEMES BY INTERFERENCE RISK FOR LATIN-SCRIPT USERS

Group	*N*	*Graphemes*	*Description*
F - familiar	10	А, Ә, Б, Е, К, М, О, С, Т, Һ	Visually and phonemically similar to Latin counterparts. Low risk.
D - deceptive	11	В, Ө, Қ, Н, Ң, Р, У, Ұ, Ү, Х, І	Visually similar to Latin letters but different phonemic values. High risk.
U - unfamiliar	21	Г, Ғ, Д, Ж, З, И, Й, Л, П, Ф, Ш, Ы, Ю, Я, Ч, Щ, Э, Ё, Ъ, Ь, Ц	No visual equivalent in Latin. Learned through memorization.

The deceptive group represents the highest methodological risk because visual familiarity may automatically activate an incorrect L1-based sound value.

Table 4 shows that half of the alphabet is visually accessible to Latin-script users, but only 10 graphemes allow safe direct transfer. The 11 deceptive graphemes make visual analogy risky because a familiar shape does not guarantee the correct sound value. The unfamiliar group requires memorization, but it may be less misleading because it does not invite a false L1-based reading.

B. Coefficients of Visual Familiarity, Direct Transfer, and Interference Risk

Four coefficients were calculated from the classification in Table 4 to transform descriptive alphabetic difficulty into operational indicators.

© 2026 ACADEMY PUBLICATION

TABLE 5
QUANTITATIVE COEFFICIENTS OF GRAPHEME-PHONOLOGICAL RISK

Coefficient	*Formula*	*Calculation*	*Value*	*Interpretation for Pre-A1 instruction*
VFC Visual Familiarity Coefficient	(F + D) / N	(10 + 11) / 42	0.50	Moderate visual familiarity: half of the alphabet looks at least partly accessible to Latin-script users.
DTC Direct Transfer Coefficient	F / N	10 / 42	0.238	Low positive transfer: fewer than one quarter of the graphemes can be treated as relatively safe.
MFR Misfire Rate	D / (F + D)	11 / 21	0.524	High misidentification risk among visually familiar signs: deceptive forms slightly outnumber safe familiar forms.
IRC Interference Risk Coefficient	(2D + U) / 2N	(2 × 11 + 21) / 84	0.512	Overall high interference risk: explicit script instruction is required before full A1 progression.

TABLE 6
ROBUSTNESS AND SENSITIVITY CHECKS FOR MFR AND IRC

Scenario	**F / D / U**	**MFR**	**IRC**	**Change from baseline**	**Interpretation**
Baseline classification	10 / 11 / 21	0.524	0.512	—	High misfire and overall interference risk.
One familiar grapheme reclassified as deceptive	9 / 12 / 21	0.571	0.536	+0.048 / +0.024	The risk increases but remains in the same high-risk zone.
One deceptive grapheme reclassified as familiar	11 / 10 / 21	0.476	0.488	-0.048 / -0.024	The indices decrease slightly, but the general interpretation remains stable.
One unfamiliar grapheme reclassified as deceptive	10 / 12 / 20	0.545	0.524	+0.022 / +0.012	A small increase confirms sensitivity to deceptive visual similarity.
One deceptive grapheme reclassified as unfamiliar	10 / 10 / 22	0.500	0.500	-0.024 / -0.012	The indices remain near the baseline high-risk threshold.

Table 6 adds a robustness check for MFR and IRC by testing one-grapheme reclassification scenarios. The strongest change is limited to 0.048 for MFR and 0.024 for IRC, while all scenarios keep the interpretation within the same methodological range: visually deceptive graphemes remain the main source of alphabet-level interference. Therefore, the coefficients should be understood as stable descriptive risk indicators for Pre-A1 course design rather than as population estimates requiring inferential statistics.

The coefficients in Table 5 show moderate visual familiarity, low direct transfer, high misfire risk among familiar-looking symbols, and a substantial overall interference risk.

The key result is the gap between VFC and DTC. Visual accessibility covers half of the alphabet, but safe transfer covers only about one quarter. Thus, an analogy-based strategy is unreliable at Pre-A1 and must be replaced by explicit contrastive instruction. This also explains why the alphabet should not be presented as a short technical introduction before grammar.

C. Recurrent Grapheme-Phonological Error Patterns in Classroom Data

Structured observation showed that alphabet-level risk appeared in recurring learner productions, especially when a deceptive or unfamiliar grapheme was connected with a Kazakh-specific phoneme.

© 2026 ACADEMY PUBLICATION

TABLE 7
RECURRENT GRAPHEME-PHONOLOGICAL ERROR PATTERNS DOCUMENTED IN THE OXFORD COURSE

Diagnostic focus	*Target form / category*	*Typical learner production or error*	*Linguistic context*	*Evidence available in the manuscript materials*
Uvular stop	Kazakh /q/; grapheme Қ; example: qalada (“in the city”)	/q/ → /k/	Initial sound-letter mapping and basic lexical reading	Reported as one of the stable substitutions; belongs to the high-risk Kazakh-specific sound group.
Voiced uvular fricative	Kazakh /ʁ/; grapheme Ғ; example: ğalym (“scholar”)	/ʁ/ → /g/	Reading and repetition of forms containing Kazakh-specific consonants	Reported as a predictable substitution based on the closest L1 category.
Front rounded vowel	Kazakh /ø/; grapheme Ө; example: özen (“river”)	/ø/ → /y/ or another English-based vowel approximation	Perception and production of vowels absent from English	Reported as a stable substitution in the discussion of Oxford data.
Back unrounded vowel	Kazakh /ɯ/; grapheme Ы; example: balyqşy (“fisherman”)	Unstable English-based vowel approximation	Vowel classification before suffix selection	Included among the Kazakh sounds that produced stable substitution difficulty.
Control case	Kazakh /ŋ/; grapheme Ң; example: jaŋa (“new”)	No stable difficulty reported due to English ng analogy	Sound-letter mapping and syllable-final/nasal recognition	Presented as an internal control example: L1 analogue reduces the problem.

Table 7 supports the phonological claims by linking target forms, typical learner productions, linguistic context, and diagnostic interpretation. The substitutions /q/ → /k/, /ʁ/ → /g/, and /ø/ → /y/ recurred across classroom data and correspond to predictable assimilation to English-like categories. By contrast, /ŋ/ was easier because English already provides a comparable category through the English sequence ng in words such as sing and running, which functions as an internal control example. The table therefore shows not only that some sounds were difficult, but also that difficulty depended on the availability or absence of a usable L1 perceptual category. These data support a diagnostic chain: a deceptive or unfamiliar grapheme can trigger an incorrect sound value; the unstable category can affect vowel classification; and this can lead to suffix-selection errors.

D. Vowel Harmony and Suffix-Vowel Selection

After the grapheme-phonological barrier, learners faced morphonological difficulties connected with vowel harmony. The back/front distinction was learned more effectively when explained as a vowel-class system rather than through the potentially misleading hard/soft terminology.

The central result is that vowel harmony directly influenced early grammar. To choose a suffix, learners had to identify the stem vowel class before selecting a back or front variant.

TABLE 8
TWO TYPES OF SUFFIX VOWELS IN EARLY KAZAKH MORPHONOLOGY

Type	*Back-vowel variants*	*Front-vowel variants*	*Examples in Latin transliteration with translation*
a / e	-da, -ta, -lar, -dar, -ğa	-de, -te, -ler, -der, -ge	adam-da (“in/at a person”); adam-dar (“people”); äje-de (“at grandmother”); äje-ler (“grandmothers”)
y / i	-dy, -ty, -nyŋ, -dyŋ	-di, -ti, -niŋ, -diŋ	auyl-dy (“the village”, accusative); halyq-tyŋ (“of the people”); til-di (“the language”, accusative); mektep-tiŋ (“of the school”)

Table 8 shows that even elementary suffixes require the learner to connect stem vowel quality with suffix choice. This makes vowel harmony an operational grammar skill, not only a pronunciation topic. At Pre-A1, learners need repeated exposure to transparent suffix families before they are expected to manipulate more complex paradigms.

E. Consonant Assimilation and Affixal Allomorphy

Learners also encountered consonant-conditioned allomorphy, especially in the dative case and predicative endings. They often expected one invariant suffix but met variants such as -ğa/-ge/-qa/-ke, -a/-e after first- and second-person possessive markers, and -na/-ne after third-person possessives. A comparable issue appeared in *men mūğalimmın* (“I am a teacher”) versus *biz mūğalimbiz* (“we are teachers”), where the expected symmetrical form *biz mūğalimmiz* is incorrect. This shows that affix selection depends on several parameters, including vowel class, consonant type, and morphological position.

Table 9 separates vowel-driven and consonant-driven sources of morphonological error.

© 2026 ACADEMY PUBLICATION

TABLE 9
DIAGNOSTIC SEPARATION OF EARLY MORPHONOLOGICAL ERROR SOURCES

Error source	*Observable learner difficulty*	*Representative Kazakh example*	*Diagnostic interpretation*	*Instructional response*
Vowel-harmony source	Learner does not reliably distinguish back/front vowel class of the stem.	änşi (“singer”), not *änşy; balyqşy (“fisherman”), not *balyqşi	Suffix choice fails because harmonic class is unstable.	Perceptual training for vowel classes before paradigm expansion.
Consonant-assimilation source	Learner expects one dative suffix instead of the variants -ğa/-ge/-qa/-ke.	qala-ğa (“to the city”); mektep-ke (“to the school”)	Suffix choice fails because final consonant and voicing conditions are not yet operational.	Contrastive drills organized by stem-final consonant type.
Possessive-position source	Learner applies the general dative pattern after possessive forms.	ata-m-a (“to my father”); bala-sy-na (“to his/her child”)	Morphological position changes the allomorph set.	Teach possessive + case combinations as a separate micro-pattern.
Predicative-ending source	Learner creates a symmetrical form by analogy.	men mūğalimmın (“I am a teacher”) vs biz mūğalimbiz (“we are teachers”); not *biz mūğalimmiz	Analogy is overgeneralized from one person/number form to another.	Introduce personal endings through controlled contrasts, not as a single invariant template.

Table 9 shows that affixal variation must be treated diagnostically. The teacher should determine whether the error comes from vowel classification, consonant assimilation, possessive morphology, or false analogy between personal endings. This separation is pedagogically important because the same incorrect suffix may require different types of feedback.

F. Results-Based Justification of an Independent Pre-A1 Module

Taken together, the results justify an independent Pre-A1 stage for English-speaking learners of Kazakh. VFC, DTC, MFR, and IRC identify alphabetic risk; classroom data identify recurring substitutions; and the morphonological analysis shows that early pronunciation problems affect grammar. The evidence therefore supports a sequential model in which script, sound, vowel harmony, and allomorphy are treated as connected parts of one diagnostic zone.

The initial course should therefore not move directly from alphabet presentation to ordinary A1 grammar. A separate Pre-A1 module should establish script recognition, sound-letter mapping, perceptual contrasts, vowel-harmony recognition, controlled suffix variation, and embedded correction before grammatical complexity increases. The results support a sequence in which the learner first stabilizes the graphic and phonological basis of the word, then uses it for controlled morphonological choice.

V. DISCUSSION

The study aimed to justify an independent Pre-A1 stage for English-speaking learners of Kazakh by identifying grapheme-phonological and morphonological difficulties. The results show that early difficulty is not limited to alphabet memorization or isolated pronunciation, but forms a chain from script recognition to phoneme categorization, vowel-harmony recognition, and suffix selection. The combination of structured observation, written assignments, audio-supported exercises, and coefficient-based modeling makes this chain visible and pedagogically interpretable.

These findings confirm research showing that orthographic input influences L2 phonological acquisition and word learning (Bassetti, 2006; Hayes-Harb & Barrios, 2021). They also extend work on Russian Cyrillic, where familiar-looking letters create phono-lexical errors, by showing that Kazakh-specific graphemes add morphonological consequences (Showalter, 2020; Shepperd, 2024). The contrast between visual familiarity and safe transfer supports the diagnostic value of MFR and IRC, because the most dangerous cases are not necessarily the most visually unfamiliar ones. In this respect, the study refines the general idea of cross-script difficulty: error risk depends on the relationship between visual familiarity, phonological equivalence, and the grammatical role of the sound involved.

The phonological results align with SLM-r and PAM-L2: sounds without stable English equivalents, including /q/, /ʁ/, /ɯ/, /ø/, and /ʏ/, tend to be assimilated to existing L1 categories (Best & Tyler, 2007; Flege & Bohn, 2021). The data add that a written symbol may trigger this assimilation before auditory processing is stabilized. Vowel-harmony errors further show that perception failure and rule failure must be separated in feedback: a wrong suffix may indicate a misunderstood rule, but it may also indicate an unstable vowel category. This distinction is important because repeating a grammatical explanation will not solve an error caused by inaccurate perception of the stem vowel.

The scientific contribution is the integration of orthographic transfer, L2 phoneme perception, and Kazakh morphonology into one Pre-A1 diagnostic model. Practically, the study supports a short but systematic module with script orientation, contrastive sound-letter mapping, perceptual training, vowel-harmony recognition, controlled affixal variation, and embedded correction. Such sequencing can reduce decoding load and make later A1 instruction more stable. It also helps teachers decide whether a learner’s error requires phonetic training, grapheme contrast, or morphonological rule practice. The proposed model is especially relevant for intensive courses, where the pressure to move quickly to vocabulary and grammar can hide unresolved pre-threshold difficulties.

Future research should move from exploratory diagnosis to broader validation. The proposed model should be tested with larger samples, in other institutions, and with learners from different L1 backgrounds. Instrumental acoustic analysis,

© 2026 ACADEMY PUBLICATION

including formant measurements for vowel contrasts and spectral or temporal parameters for place contrasts, would strengthen the validation of instructor IPA transcription. Comparative work with other Turkic languages would show whether VFC, DTC, MFR, and IRC can function as broader tools for diagnosing cross-script and morphonological risk. Longitudinal studies could also determine whether explicit Pre-A1 instruction reduces later A1 errors in suffix selection and pronunciation. Such research would help refine the boundary between language-specific beginner difficulty and transferable principles of cross-script pedagogy.

VI. Conclusion

The study achieved its aim by showing that English-speaking learners face a systematic pre-threshold difficulty in Kazakh. Structured observation, written assignments, audio-supported phonetic exercises, and grapheme classification demonstrated that Cyrillic recognition, sound-letter mapping, phoneme perception, vowel harmony, and affixal allomorphy form an interdependent learning chain rather than separate beginner topics. This confirms the need to treat Pre-A1 as a diagnostic stage with its own content and sequence.

The main contribution is the integration of classroom observation with a formalized model of grapheme-phonological risk. By distinguishing familiar, deceptive, and unfamiliar graphemes and calculating VFC, DTC, MFR, and IRC, the article makes the Pre-A1 stage diagnostically visible and methodologically justified. It also specifies OOPHI as a mechanism through which written input can trigger sound-category misclassification. The practical outcome is a structured basis for designing a Kazakh Pre-A1 module before full A1 progression.

The limitations are the small sample, the single-instructor observation format, and the absence of acoustic measurements. Nevertheless, the Oxford case reveals stable mechanisms that can inform course design. Further studies should test the model on larger cohorts, add instrumental phonetic validation, and compare Kazakh with other Turkic languages and non-Latin writing systems. Such work would clarify which elements of the model are language-specific and which can be transferred to broader cross-script language pedagogy.

Declaration of the Use of AI

The authors used AI-assisted tools to support translation of the manuscript from Russian into English, followed by human revision and verification. Claude Opus 4.6 was used for language assistance only.

Acknowledgements

This article was developed within the framework of the program-targeted funding project BR24993244 (2024–2026) “Improving the National Corpora of the Kazakh Language as a basis for the Smart Texts megaproject and Kazakh-language artificial intelligence, and developing its internal corpora”.

References

[1] Asselborn, T., Johal, W., Tleubayev, B., Zhexenova, Z., Dillenbourg, P., McBride, C., & Sandygulova, A. (2021). The transferability of handwriting skills: From the Cyrillic to the Latin alphabet. *NPJ Science of Learning*, *6*, Article 6. https://doi.org/10.1038/s41539-021-00084-w

[2] Ayapova, T. T., & Thomas, A. D. (1992). *Learn the Kazakh language in 70 steps: Using 200 sentence models*. Dunwoody Press.

[3] Bakytzhanova, A., Bekmukhambetova, S. B., Utemissova, E. K., Zhanzhigitov, S., Alpysbayeva, A. Y., Omarbayeva, G. S., & Tatiyeva, S. S. (2025). Linguo-conceptual and linguocultural analysis of the concept “memory”: Semantic typology of Kazakh-Russian-English phraseology. *Theory and Practice in Language Studies*, *15*(12), 4162–4171. https://doi.org/10.17507/tpls.1512.36

[4] Bassetti, B. (2006). Orthographic input and phonological representations in learners of Chinese as a foreign language. *Written Language & Literacy*, *9*(1), 95–114. https://doi.org/10.1075/wll.9.1.07bas

[5] Batayeva, Z. (2012). *Colloquial Kazakh: The complete course for beginners*. Routledge.

[6] Bazarbayeva, Z. M. (2012). *Fundamentals of Kazakh phonology*. Institute of State Language Development.

[7] Best, C. T., & Tyler, M. D. (2007). Nonnative and second-language speech perception. In O.-S. Bohn & M. J. Munro (Eds.), *Language experience in second language speech learning* (pp. 13–34). John Benjamins.

[8] Brennan, C., & Kiskin, J. (2020). Distinct benefits given large versus small grain orthographic instruction for English-speaking adults learning to read Russian Cyrillic. *Journal of Psycholinguistic Research*, *49*(6), 915–933. https://doi.org/10.1007/s10936-019-09684-5

[9] Cook, V. J., & Bassetti, B. (Eds.). (2005). *Second language writing systems*. Multilingual Matters.

[10] Council of Europe. (2020). *Common European Framework of Reference for Languages: Learning, teaching, assessment—Companion volume*. Council of Europe Publishing.

[11] Derwing, T. M., & Munro, M. J. (2015). *Pronunciation fundamentals: Evidence-based perspectives for L2 teaching and research*. John Benjamins.

[12] Erdener, V. D., & Burnham, D. K. (2005). The role of audiovisual speech and orthographic information in nonnative speech production. *Language Learning*, *55*(2), 191–228. https://doi.org/10.1111/j.0023-8333.2005.00303.x

[13] Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 233–277). York Press.

© 2026 ACADEMY PUBLICATION

[14] Flege, J. E., & Bohn, O.-S. (2021). The revised Speech Learning Model (SLM-r). In R. Wayland (Ed.), *Second language speech learning: Theoretical and empirical progress* (pp. 3–83). Cambridge University Press.
[15] Hayes-Harb, R., & Barrios, S. (2021). The influence of orthography in second language phonological acquisition. *Language Teaching*, *54*(3), 297–326. https://doi.org/10.1017/S0261444820000658
[16] Hayes-Harb, R., Nicol, J., & Barker, J. (2010). Learning the phonological forms of new words: Effects of orthographic and auditory input. *Language and Speech*, *53*(3), 367–381. https://doi.org/10.1177/0023830910371460
[17] Jaccard, P. (1901). Étude comparative de la distribution florale dans une portion des Alpes et des Jura [Comparative study of floral distribution in a portion of the Alps and the Jura]. *Bulletin de la Société Vaudoise des Sciences Naturelles*, *37*, 547–579.
[18] Krippes, K. A. (1996). *Kazakh grammar with affix list*. Dunwoody Press.
[19] Kucherbayeva, D., & Smagulova, J. (2023). Language revitalization: Challenges for Kazakh in higher education. *Journal of Eurasian Studies*, *14*(2), 166–178. https://doi.org/10.1177/18793665231189326
[20] Lee, J., Jang, J., & Plonsky, L. (2015). The effectiveness of second language pronunciation instruction: A meta-analysis. *Applied Linguistics*, *36*(3), 345–366. https://doi.org/10.1093/applin/amu040
[21] McCollum, A. G. (2018). Vowel dispersion and Kazakh labial harmony. *Phonology*, *35*(2), 287–326. https://doi.org/10.1017/S0952675718000052
[22] McCollum, A. G. (2020). Vowel harmony and positional variation in Kyrgyz. *Laboratory Phonology: Journal of the Association for Laboratory Phonology*, *11*(1), Article 25. https://doi.org/10.5334/labphon.247
[23] Muhamedowa, R. (2016). *Kazakh: A comprehensive grammar*. Routledge.
[24] Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge University Press.
[25] Ormanova, A. B., Anafinova, M. L., Ospanova, D. Zh., & Tleshova, Z. K. (2025). A corpus-based approach in vocabulary research: Defining the word of the year 2023 in Kazakh. *Theory and Practice in Language Studies*, *15*(3), 677–687. https://doi.org/10.17507/tpls.1503.02
[26] Ringbom, H., & Jarvis, S. (2011). The importance of cross-linguistic similarity in foreign language learning. In M. H. Long & C. J. Doughty (Eds.), *The handbook of language teaching* (pp. 106–118). Wiley-Blackwell.
[27] Rsaliyeva, N. M., & Chukayeva, T. K. (2025). Corpus-based approach to teaching the Kazakh language: Adapted experience for anglophones. *Tiltanym*, *100*(4), 93–111. https://doi.org/10.55491/2411-6076-2025-4-93-111
[28] Saito, K. (2012). Effects of instruction on L2 pronunciation development: A synthesis of 15 quasi-experimental intervention studies. *TESOL Quarterly*, *46*(4), 842–854. https://doi.org/10.1002/tesq.67
[29] Saito, K., & Plonsky, L. (2019). Effects of second language pronunciation teaching revisited: A proposed measurement framework and meta-analysis. *Language Learning*, *69*(3), 652–708. https://doi.org/10.1111/lang.12345
[30] Sarmurzin, Y., Amanzhol, N., Toleubayeva, K., Zhunusova, M., Amanova, A., & Abiyr, A. (2024). Challenging aspects of Kazakhstan's trilingual education policy: Evidence from a literature review. *Asia Pacific Education Review*, *25*(4), 801–811. https://doi.org/10.1007/s12564-023-09823-7
[31] Shepperd, L. (2024). Cross-scriptal orthographic influence on second language phonology. *Languages*, *9*(6), Article 210. https://doi.org/10.3390/languages9060210
[32] Showalter, C. E. (2018). Impact of Cyrillic on native English speakers' phono-lexical acquisition of Russian. *Language and Speech*, *61*(4), 565–576. https://doi.org/10.1177/0023830918761489
[33] Showalter, C. E. (2020). Russian phono-lexical acquisition and orthographic input. *Studies in Second Language Acquisition*, *42*(2), 255–277. https://doi.org/10.1017/S0272263119000585
[34] Simon, E., Chambless, D., & Kickhöfel Alves, U. (2010). Understanding the role of orthography in the acquisition of a non-native vowel contrast. *Language Sciences*, *32*(3), 380–394. https://doi.org/10.1016/j.langsci.2009.07.001
[35] Thomson, R. I., & Derwing, T. M. (2015). The effectiveness of L2 pronunciation instruction: A narrative review. *Applied Linguistics*, *36*(3), 326–344. https://doi.org/10.1093/applin/amu076
[36] Webb, S., & Nation, P. (2017). *How vocabulary is learned*. Oxford University Press.
[37] Yelubayeva, P., Khamidova, A., Berkinbayeva, G., & Avakova, R. (2025). Addressing language education challenges in Kazakhstan for sustainable development. *European Journal of Language Policy*, *17*(1), 65–96. https://doi.org/10.3828/ejlp.2025.5
[38] Yin, R. K. (2018). *Case study research and applications: Design and methods* (6th ed.). SAGE.
[39] Zhanzhigitov, S., Abdualiuly, B., Doszhan, R., Serikbayeva, Z., Aigul, M., & Balzhan, Y. (2022). The problem of systematization of terms in Turkic languages and technologies of teaching at universities. *Cypriot Journal of Educational Science*, *17*(4), 1077–1089. https://doi.org/10.18844/cjes.v17i4.7100
[40] Zhumashova, Z., Aksholakova, A., Omirbekova, A., & Kulbekova, B. (2023). Challenges and recommendations for teaching and learning Kazakh as a second language. *International Journal of Society, Culture & Language*, *11*(2), 119–131. https://doi.org/10.22034/ijscl.2023.1995923.2967
[41] Zhunisbek, A. (2018). *Qazaq til biliminin maseleleri* [Issues of Kazakh linguistics]. Abzal-Ai.

Anar Fazylzhan is a Candidate of Philology and Associate Professor at A. Baitursynuly Institute of Linguistics, Almaty, Kazakhstan. She works in the field of Kazakh linguistics and has held research and academic positions connected with the study and teaching of the Kazakh language. Her scholarly interests include Kazakh phonetics, intonation, morphology, syntax, paralinguistics, computer linguistics, language standardization, and the development of educational and reference materials. She has contributed to Kazakh language textbooks, orthographic and orthoepic reference resources, bilingual and normative dictionaries, and scientific publications on Kazakh philology. In the present article, her expertise is connected with the phonetic, phonological, and methodological interpretation of Kazakh language material. Email: anar.fazylzhan@tilbilimi.kz ORCID iD: https://orcid.org/0000-0002-0562-4846

© 2026 ACADEMY PUBLICATION

Didar Sadyk Aidaruly (corresponding author) is a PhD doctoral student at Al-Farabi Kazakh National University and a Research Fellow at A. Baitursynuly Institute of Linguistics, Almaty, Kazakhstan. He holds a Master's degree and works in the field of Kazakh as a second and foreign language. His academic and professional experience is related to Kazakh phonetics, phonological processes, linguodidactics, grapheme-phonetic and phonological interference, and the methodology of teaching Kazakh to English-speaking learners. He participates in academic projects aimed at teaching and promoting Kazakh in international educational contexts, including work connected with Kazakh language instruction at the University of Oxford. His current research interests include Kazakh phonetics and phonology, second language acquisition, corpus and digital technologies in language teaching, and the development of teaching models for foreign learners of Kazakh. Email: didarsadyk2@gmail.com ORCID iD: https://orcid.org/0000-0002-8271-7165

Madina Abayeva is a Candidate of Philology and Associate Professor at A. Baitursynuly Institute of Linguistics, Almaty, Kazakhstan. Her scholarly work is connected with Kazakh philology, applied linguistic description, and the analysis of language material relevant to contemporary Kazakh linguistics. Her research interests include problems of Kazakh language structure, linguistic description, phonetic and phonological issues, and the methodological interpretation of Kazakh language data. In the present study, her contribution is associated with the linguistic and methodological analysis of Kazakh as a foreign language and the interpretation of grapheme-phonological and morphonological difficulties. Her academic profile supports the broader aim of adapting Kazakh linguistic knowledge to educational and international research contexts. Email: madina.abaeva@gmail.com ORCID iD: https://orcid.org/0000-0001-6463-9354

Yelimay Shynatay is a PhD doctoral student at Al-Farabi Kazakh National University and a Research Fellow at A. Baitursynuly Institute of Linguistics, Almaty, Kazakhstan. He holds a Master's degree and specializes in linguistics, with academic training connected with Kazakh philology and language theory. His research interests include linguistics, psycholinguistics, second language acquisition, and the methodology of teaching foreign and second languages. His work is relevant to the present article because the study focuses on the initial acquisition of Kazakh by English-speaking learners and on the interaction between script, phonology, and morphonology. He is engaged in research on language learning processes and methodological issues of foreign-language instruction. Email: yelimayshynatay@gmail.com ORCID iD: https://orcid.org/0000-0002-7554-8949

© 2026 ACADEMY PUBLICATION