ACADEMIA ONE
PEER REVIEWED JOURNALS

Open Academia: Journal of Scholarly Research
Volume 4, Issue 01 January, 2026
ISSN (E): 2810-6377
Website: https://academiaone.org/index.php/4

Paroemiological Units In Cognitive And Linguocultural Research

Javod Aminbaevich Kurbanbaev
Dean of the Faculty of Foreign Languages, PhD (Pedagogy), Associate Professor
Berdakh Karakalpak State University, Nukus, Uzbekistan

Abstract. The article discusses how AI-assisted tools can be integrated into linguocultural vocabulary instruction in Uzbek universities without replacing linguistic evidence or teacher judgment. It is argued that culturally marked Russian vocabulary (realia, cultural keywords, precedent-related units, and pragmatically sensitive expressions) requires learning environments where students can observe authentic usage patterns, interpret cultural scripts, and rehearse genre-appropriate choices. The paper proposes an instructional model that combines corpus-informed noticing with AI-supported practice: micro-corpora provide empirical contexts and collocational profiles, while AI-based dialogic simulation supports scenario rehearsal for politeness, stance, and institutional communication. The article outlines design principles for pedagogical control, including transparency of evidence, constraint-based prompting, teacher validation, and ethical safeguards (privacy, bias awareness, and academic integrity). The conclusion emphasizes that AI is most effective when positioned as a structured practice layer built on corpus evidence and pragmatics-focused rubrics, leading to measurable gains in reading interpretation, register control, and communicative appropriateness.
Keywords: AI-assisted learning, linguoculturology, culturally marked vocabulary, corpus-based instruction, pragmatics, register, scenario simulation, Russian as a foreign language, bilingual education, Uzbekistan

In Uzbek higher education, Russian-language proficiency is increasingly evaluated through performance in academic writing, institutional communication, and professional interaction. Yet vocabulary instruction and assessment often remain centered on translation, synonym lists, and definition recall. Such methods support basic semantic control but do not fully address culturally marked vocabulary, where meaning depends on genre conventions, evaluative stance, and culturally conventional scripts (Maslova, 2001; Kramsch, 1998). For this domain, the key pedagogical problem is not "unknown words" but "known words used in the wrong way": learners reproduce a lexical item correctly while violating pragmatic boundaries or register expectations.

AI-assisted tools have recently become attractive for language education because they can generate practice opportunities, simulate dialogs, and provide rapid feedback. However, for linguocultural vocabulary, the central risk is methodological: if AI outputs become the primary source of "meaning," learners may internalize plausible but unverified explanations, or adopt non-target-like patterns. For that reason, AI must be positioned as a controlled practice layer grounded in empirical evidence, not as a replacement for corpus contexts, dictionaries, and teacher interpretation. Corpus linguistics remains essential because it shows how words behave in authentic discourse through repeated contexts and collocations, which are often decisive for cultural and pragmatic meaning (Sinclair, 1991; McEnery & Hardie, 2012).

A workable model is a two-layer architecture for instruction. The first layer is **evidence**, provided by micro-corpora curated from course texts and reliable sources. The second layer is **practice**, supported by AI-driven interaction that is constrained by the evidence layer. In the evidence layer, teachers compile small topic-based corpora (e.g., education, social issues, university life) and extract typical contexts for target items. This aligns with data-driven

ACADEMIA ONE
PEER REVIEWED JOURNALS
Open Academia: Journal of Scholarly Research
Volume 4, Issue 01 January, 2026
ISSN (E): 2810-6377
Website: https://academiaone.org/index.php/4

learning principles, where learners develop sensitivity to usage patterns through guided observation (Boulton, 2010; O'Keeffe, McCarthy, & Carter, 2007). In the practice layer, AI is used to generate controlled tasks: dialog prompts, paraphrase exercises, register-shift rewrites, and scenario simulations that require learners to apply observed patterns under communicative constraints.

The instructional sequence can be implemented through four steps. First, teachers select culturally marked items that are recurrent and communicatively risky in Uzbek classrooms: realia and institution-related vocabulary, cultural keywords, precedent-related expressions, and politeness-sensitive formulas. Second, each item is profiled using micro-corpus evidence: typical collocations, genre distribution, and evaluative tendencies are summarized in a compact entry. Third, AI-supported tasks are built with constraint-based prompting: the prompt explicitly restricts the AI to the micro-corpus patterns and to the target genre, and it requires the AI to cite which collocations or contextual cues justify an output. This "traceability" principle makes practice more accountable and prevents learners from treating AI production as unquestionable authority. Fourth, learner performance is assessed with analytic rubrics that separate semantic adequacy from register alignment, collocational naturalness, pragmatic effect, and cultural-script fit (Bachman & Palmer, 1996).

Within this model, AI is particularly useful for pragmatic rehearsal. Pragmatics is difficult to teach through static lists because it depends on social roles, imposition, distance, and discourse goals. Scenario-based tasks can operationalize these conditions: the learner must choose lexical items appropriate for writing to a professor, responding to administrative feedback, disagreeing politely in a seminar, or commenting on a news excerpt in an academically acceptable tone. Pragmatics research supports scenario-based instruction, especially when it includes explicit reflection on linguistic choices and their social effects (Ishihara & Cohen, 2010; Taguchi, 2015). AI can provide multiple scenario variants quickly, which helps learners practice transfer across contexts rather than memorize a single template.

At the same time, linguocultural learning requires careful handling of cultural scripts. Cultural scripts should not be presented as rigid stereotypes; they are tendencies supported by conventional usage and discourse expectations (Sharifian, 2017). AI-generated cultural explanations can easily become overgeneralized. Therefore, the teacher's role is to keep scripts bounded: each script should be linked to observable contexts and marked as genre-sensitive. A practical classroom method is to pair every AI-supported scenario with corpus lines that show how the target word functions in similar communicative situations. Learners are then asked to justify their lexical choice by pointing to evidence. This approach maintains academic rigor while still benefiting from AI's capacity for interactive practice.

Another domain where AI can add value is register and genre transformation. Uzbek learners often face "register drift," where informal or media-styled vocabulary appears in academic writing. AI can generate controlled rewriting tasks: a conversational message is rewritten into institutional email format, or a journalistic paragraph is transformed into neutral academic prose. However, to avoid mechanical rewriting, prompts must require learners to explain why specific words were replaced (e.g., evaluation too strong, stance too categorical, collocation non-academic). Such reflection is consistent with discourse-oriented views of academic writing, where lexical choices signal stance and credibility (Hyland, 2005).

Ethical and methodological safeguards are necessary. First, privacy and data protection must be observed: learner texts used for micro-corpora should be anonymized, and institutional messages should be handled cautiously. Second, bias awareness matters: AI-generated dialogs can implicitly reproduce stereotypes, which is particularly harmful in intercultural teaching. Third, academic integrity must be protected: AI should be framed as a practice environment, not as a tool for producing graded submissions. Education research on AI adoption highlights

the importance of governance and teacher mediation rather than unregulated use (Zawacki-Richter, Marín, Bond, & Gouverneur, 2019). In practical terms, teachers can require process evidence (draft history, justification notes, corpus references) to ensure that learning remains learner-authored.

The effectiveness of AI-assisted linguocultural modules should be evaluated with criteria aligned to communicative outcomes. Translation accuracy alone is insufficient; instead, assessment should measure whether learners can (a) interpret culturally implied evaluation in reading, (b) select register-appropriate vocabulary, (c) maintain collocational stability, (d) achieve pragmatic appropriateness, and (e) avoid false equivalence by articulating script boundaries. Analytic rubrics are suitable for this purpose because they make progress visible and support formative feedback (Bachman & Palmer, 1996). When combined with tagged feedback banks, teachers can accumulate a local evidence base of recurring errors and successful strategies, gradually improving both instruction and assessment.

In conclusion, AI-assisted tools can strengthen linguocultural vocabulary learning in Uzbek universities when they are integrated into a corpus-based and pragmatically controlled design. Micro-corpora provide the empirical foundation for meaning-in-use, while AI supports repeated scenario rehearsal, genre transformation, and reflective feedback. The key condition is methodological discipline: AI outputs must remain traceable to evidence, constrained by genre requirements, and validated through teacher judgment and analytic rubrics. Under these conditions, AI becomes not a shortcut for vocabulary learning, but an instrument for scaling high-quality practice and improving intercultural communicative competence.

References

1. Bachman, L. F., & Palmer, A. S. (1996). *Language testing in practice: Designing and developing useful language tests*. Oxford University Press.
2. Boulton, A. (2010). Data-driven learning: Taking the computer out of the equation. *Language Learning, 60*(3), 534–572.
3. Hyland, K. (2005). *Metadiscourse: Exploring interaction in writing*. Continuum.
4. Ishihara, N., & Cohen, A. D. (2010). *Teaching and learning pragmatics: Where language and culture meet*. Pearson.
5. Kramsch, C. (1998). *Language and culture*. Oxford University Press.
6. Maslova, V. A. (2001). *Lingvokul'turologiya* [Linguoculturology]. Akademiya.
7. McEnery, T., & Hardie, A. (2012). *Corpus linguistics: Method, theory and practice*. Cambridge University Press.
8. O'Keeffe, A., McCarthy, M., & Carter, R. (2007). *From corpus to classroom: Language use and language teaching*. Cambridge University Press.
9. Sharifian, F. (2017). *Cultural linguistics: Cultural conceptualisations and language*. John Benjamins.
10. Sinclair, J. (1991). *Corpus, concordance, collocation*. Oxford University Press.
11. Taguchi, N. (2015). Instructed pragmatics at a glance: Where instructional studies were, are, and should be going. *State-of-the-art article*. *Language Teaching, 48*(1), 1–50.
12. Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education – Where are the educators? *International Journal of Educational Technology in Higher Education, 16*(39), 1–27.