Information about the authors:

Ayazhan Boltbergenova* – master's student of Shakarim University, Republic of Kazakhstan, Semey, e-mail: aazan_01@mail.ru, ORCID: https://orcid.org/0009-0002-0475-964X.

Baktygul Abdykhanova – PhD, Department of foreign languages high school of philology, Shakarim University, Republic of Kazakhstan, Semey, e-mail: Abdychanovab@mail.ru, ORCID: https://orcid.org/0000-0001-8076-2336.

Авторлар туралы мәліметтер:

Болатбергенова Аяжан Дарынқызы* – Шәкәрім университетінің магистранты, Қазақстан Республикасы, Семей, e-mail: aazan_01@mail.ru, ORCID: https://orcid.org/0009-0002-0475-964X.

Абдыханова Бактыгуль Айдапкелевна – PhD, шет тілдер кафедрасының меңгерушісі, Шәкәрім университеті, Қазақстан Республикасы, Семей, e-mail: Abdychanovab@mail.ru, ORCID: https://orcid.org/0000-0001-8076-2336.

Сведение об авторах:

Болатбергенова Аяжан Дарынкызы* – магистрант Шәкәрім университета, Республика Казахстан, Семей, e-mail: aazan_01@mail.ru, ORCID: https://orcid.org/0009-0002-0475-964X.

Абдыханова Бактыгуль Айдапкелевна – PhD, заведующий кафедрой иностранных языков, Шәкәрім университета, Республика Казахстан, Семей, e-mail: Abdychanovab@mail.ru, ORCID: https://orcid.org/0000-0001-8076-2336.

Received 23.07.2026
Revised 19.08.2026
Accepted 27.09.2026

https://doi.org/10.53360/3080-3861-2026-3(7)-23
IRSTI: 16.31.51

S.I. Ikhsanova[1*], Zh.U. Kalimova[2]
Shakarim University,
071412, Republic of Kazakhstan, Semey, 20 A Glinka Str.
[*]ORCID: 0009-0007-5210-8194
[*]e-mail: sofyaikhsanova578@gmail.com

USING ARTIFICIAL INTELLIGENCE TOOLS FOR AUTOMATED TEXT DIFFERENTIATION IN ENGLISH LESSONS

Abstract. This article aims to develop a theoretically grounded, criteria-based framework for evaluating the quality of AI-generated text adaptations in EFL middle school settings. The study is theoretical in nature and employs a systematic critical literature review of 20 peer-reviewed sources (2023-2026), supplemented by comparative analysis and theoretical synthesis. Three analytical procedures were applied: comparative reading of selected sources, classification of research gaps, and deductive theoretical integration. The central finding is a unified three-criterion evaluation framework, comprising difficulty alignment, readability alignment, and differentiation completeness, which is derived from the integration of the Zone of Proximal Development by Vygotsky L. S., the Input Hypothesis by Krashen S. D., and the Differentiated Instruction model by Tomlinson C. A., and operationalized through CEFR level descriptors and Flesch-Kincaid readability scores. Comparative analysis of four key empirical studies confirmed that no prior work has united all three theoretical frameworks, addressed secondary school learners, or proposed a replicable evaluation methodology in this domain. The proposed framework offers EFL teachers in multilevel classrooms a principled quality assurance procedure for assessing AI-generated materials prior to classroom use, while providing a structured foundation for the empirical phase of research.

Keywords: artificial intelligence, text adaptation, differentiated instruction, digital resources, EFL reading instruction.

Introduction

Modern English classrooms include students with very different levels of language proficiency. In one class, a teacher may work with students at A1, A2, and B1 levels at the same time. This situation, commonly referred to as a mixed-proficiency classroom, presents a serious pedagogical challenge. Effective teaching requires providing each student with reading materials that match their individual level of competence. However, preparing separate text versions for every proficiency group requires time that most teachers simply do not have, given the demands of a full school workload. As a result, many students receive identical materials regardless of their ability, which significantly reduces the effectiveness of instruction.

The growing availability of generative Artificial Intelligence (AI) tools, such as ChatGPT, Grammarly, and Gemini, offers a possible and practical solution to this problem. These tools can process authentic reading texts and produce linguistically adapted versions in a very short period of time, a process commonly known as AI-powered text adaptation. Despite the rapid growth of interest in AI-assisted education, however, the existing empirical studies on AI-based text adaptation have only been conducted at the university level and in non-Central Asian contexts. No study to date has proposed a clear, evidence-based theoretical framework for evaluating how generative AI ought to adapt reading materials for different proficiency levels in middle school EFL settings. It is gap that the present article directly addresses. The aim of the present study is therefore to develop a theoretically grounded, criteria-based evaluation framework for assessing the quality of AI-generated text adaptations in EFL middle school settings, with specific reference to the Kazakhstani educational context.

The central research question of the present article is explores to what extent generative AI tools can produce linguistically accurate, CEFR-aligned adaptations of authentic texts for middle school EFL instruction, and what theoretical criteria should guide their evaluation. Since this is a theoretical study, the article does not present empirical classroom data. Instead, it conducts a systematic review of existing literature and proposes a criteria-based evaluation framework grounded in three established theories: the Zone of Proximal Development (ZPD) by Vygotsky L. S. [1], which positions AI as a digital scaffold that keeps learners within their optimal learning zone; Input Hypohesis (i+1), by Krashen S. D. [2] which defines the linguistic adjustment that adapted texts must achieve in relation to the learner's CEFR level; and the Differentiated Instruction (DI) framework by Tomlinson C. A. [3], which provides the pedagogical reason for designing level-differentiated materials in mixed-ability classrooms.

This study holds direct relevance for the educational context of Kazakhstan, where secondary schools face the specific challenge of delivering English language instruction within a trilingual policy framework that simultaneously encompasses Kazakh, Russian, and English. EFL classrooms in Kazakhstani secondary schools remain characterised by pronounced proficiency variation and significant teacher workload constraints [4; 5]. The national digital education agenda, alongside these practical constraints, makes the theoretical investigation of AI-assisted text adaptation both timely and necessary. By developing a principled evaluation framework, this article contributes both to international research in language didactics and to the applied needs of middle school educators in Central Asia.

Literature Review

The following review draws on 20 primary sources published predominantly between 2023 and 2026, selected from ERIC, Scopus, and Google Scholar based on their relevance to AI in EFL education, text adaptation, or differentiated instruction, as well as their full-text accessibility and peer-reviewed status. Sources focused solely on AI for writing production, administrative use, or non-EFL contexts were excluded. The review is organised thematically, moving from general AI integration in EFL, through the Kazakhstani context, to AI-assisted text adaptation and AI-supported differentiated instruction.

Research on Artificial Intelligence in EFL education has grown considerably in recent years, examining tools such as ChatGPT, Grammarly, and adaptive learning platforms across a variety of

instructional contexts. Talapova A. K. and Yalkonova Y. D. (2025) [6] investigated AI integration in Kazakhstani EFL settings and reported improvements in students' writing accuracy, speaking confidence, and vocabulary memory, alongside challenges related to limited digital foundation and data privacy concerns. Uteshova A. (2025) [7] reviewed global and local practices of AI integration in English language teaching, concluding that a combination of AI innovation and traditional pedagogy is essential for balanced learning outcomes. Mendekenova A. et al. (2025) [8], in a mixed-methods survey of 154 teachers from 37 countries, found AI to be the most widely adopted educational technology, rated highest for skill development, learner autonomy, and workload reduction. Laoha R., Chomthing W., and Pongpanich W. (2025) [9] identified «AI for Differentiated Instruction» as one of ten essential EFL teacher competency domains in the AI era. While this body of research confirms the broad pedagogical potential of AI in EFL classrooms, most of these studies focus on general tool adoption and teacher perceptions rather than on the specific process of producing and evaluating differentiated reading materials – the focus of the present study.

Within the Kazakhstani context specifically, Kemelbekova Z., Degtyareva X., and Seidaliyeva G. (2024) [10] examined the effect of AI chatbots on speaking skills at university level, reporting improvements in oral communicative performance, though without significant gains in pronunciation. Nogaibayeva A. and Yersultanova G. (2025) [4], through in-depth interviews with 16 secondary school EFL teachers, found pragmatic openness toward digital tools alongside concerns that AI might diminish teachers' professional role. Mektepbergenov N. et al. (2025) [11] surveyed 320 pre-service teachers and found that prior AI experience strongly predicted their intention to adopt it professionally. Sultan Y., Dautova G., and Dalle J. (2025) [12] found that among 334 philology students, AI literacy outpaced practical application skills, underscoring the need for targeted pedagogical frameworks. Fazilova A. and Kayip A. (2025) [5] identified efficiency, customisation, and innovation as benefits of AI-generated lesson plans, but noted teachers' concerns about a perceived lack of personalisation for individual student needs. None of these studies, however, addresses text adaptation or reading differentiation directly. The closest, Abildina S. et al. (2024) [13], examined differentiated instruction and digital tools for improving reading literacy, but only at primary level, leaving Grade 7-9 EFL classrooms in Kazakhstan unrepresented in this research strand. Azamatova A. et al. (2023) [14] further confirmed AI's positive effect on language achievement and motivation among Kazakhstani university students, reinforcing the broader pedagogical potential of AI-integrated instruction in the local context.

A smaller but growing body of research addresses AI-based text adaptation directly, and each study leaves a specific gap unaddressed. Murphy Odo D. (2025) [15] compared automatic text simplification software to human expert adaptation and found that automatically simplified texts improved comprehension for higher-proficiency readers but not for lower-proficiency ones – demonstrating the risk of level-undifferentiated AI output without offering criteria to prevent it. Çelik F., Yangın Ersanlı C., and Arslanbay G. (2024) [16], in an experimental study with 105 Turkish undergraduate EFL students, found that ChatGPT-simplified texts significantly improved comprehension and inference, though not reading anxiety; the study invoked the Input Hypothesis without operationalising it through CEFR descriptors or readability metrics. Aldamen H. et al. (2026) [17] similarly grounded a Jordanian university study in the Input Hypothesis and Cognitive Load Theory, showing clear benefits for lower-proficiency learners without disadvantaging higher-proficiency ones, but likewise without translating theory into replicable evaluation criteria. Nguyen T. Q. Y. (2024) [18] investigated ChatGPT's capacity for text adaptation and identified the absence of an evaluation framework as an explicit limitation, without proposing one. All four studies are set in university contexts, leaving secondary-school teachers without any principled, criteria-based framework for evaluating AI-adapted materials – the specific gap the present article addresses.

Research connecting AI to Differentiated Instruction (DI) points the same way. Rapti C. and Panagiotidis P. (2025) [19], surveying 116 foreign language educators in Greece, found that many

used AI to tailor content to student abilities but did not assess the linguistic accuracy or proficiency alignment of the output. Karataş F. (2025) [20], studying 440 teachers in Turkey, found that AI, especially ChatGPT, was used mainly to omit irrelevant content, add personalised material, and revise existing texts. Chitez M., Csürös K., and Rogobete R. (2025) [21] described a Romanian teacher-training programme in which educators came to view AI in three complementary roles: simplifier of complex texts, co-designer of instructional content, and confirmer of professional judgement. Rahmawati Y., Fitriati S., and Wahyuni S. (2025) [22] demonstrated AI-scaffolded, meaning-focused input in an Indonesian EFL flipped classroom while preserving teachers' central role in curricular decision-making. Woo D. J., Guo K., and Susanto H. (2025) [23] found that EFL students' purposes for prompting AI in story writing were often undefined or reactive, reinforcing the need for teacher-mediated, theoretically grounded frameworks. None of these studies, however, unifies the DI framework by Tomlinson C. A., the ZPD by Vygotsky L. S., and the Input Hypothesis by Krashen S. D. within a model designed to guide the evaluation of AI-adapted texts in EFL middle school settings.

Taken together, this literature points to four interconnected gaps that the present study seeks to address. Theoretically, no study to date has integrated the Zone of Proximal Development, the Input Hypothesis, and the Differentiated Instruction framework into a unified model for AI-based text evaluation, and even studies that reference the Input Hypothesis fail to implement it through CEFR descriptors. Empirically, the university-level concentration of existing text-adaptation studies leaves secondary school contexts without a criteria-based evaluation framework incorporating CEFR alignment and readability metrics such as Flesch-Kincaid scores. Contextually, the trilingual, centrally curriculated Kazakhstani middle school EFL setting – with its specific teacher workload conditions and linguistically diverse classrooms – is entirely absent from the text-adaptation literature. Methodologically, no reviewed study applies systematic literature review combined with theoretical synthesis as its primary scholarly contribution.

In light of these gaps, the present article builds most directly on the input-based studies of Aldamen H. et al. (2026) [17] and Çelik F. et al. (2024) [16] by operationalising the Input Hypothesis through explicit CEFR descriptors; responds to the call by Nguyen T. Q. Y. (2024) [18] and Murphy Odo D. (2025) [15] for a replicable, criteria-based evaluation framework; complements the Kazakhstani studies of Kemelbekova Z. et al. (2024) [10], Nogaibayeva A. and Yersultanova G. (2025) [4], and Fazilova A. and Kayip A. (2025) [5] by addressing the specific question of text quality that these works leave open; and extends the methodological approach of Chitez M. et al. (2025) [21] and Rahmawati Y. et al. (2025) [22] by grounding AI-assisted differentiation in a unified three-theory model.

Materials and Methods

The present study is theoretical in nature. It does not involve empirical data collection, classroom experiments, or direct observation of participants. Instead, the research relies on a systematic critical literature review, supported by comparative analysis and theoretical synthesis. This methodological approach is appropriate for the purposes of the study, since the central research question, which explores the extent to which generative AI tools can produce CEFR-aligned adaptations for middle school EFL instruction, requires a principled theoretical foundation before any empirical investigation can take place.

The literature search utilized three widely recognised academic databases: ERIC, Scopus, and Google Scholar, which served as the primary tertiary sources for locating relevant secondary literature. The search process employed the following keyword combinations: artificial intelligence and EFL, AI text adaptation, text simplification and language learning, differentiated instruction and AI, CEFR-aligned materials, and AI in Kazakhstan education. Three inclusion criteria guided the selection process: direct relevance to AI in EFL education, reading text adaptation, or differentiated instruction; a publication window predominantly between 2023 and 2026 to reflect current developments; and full-text accessibility in English or Russian. The study excluded sources focusing solely on AI for writing tasks, administrative functions, or non-EFL instructional contexts.

Of the 20 primary sources selected, 16 are peer-reviewed journal articles, 3 are book chapters, and 1 is a conference proceeding; 18 are in English and 2 in Russian. Geographically, the corpus includes studies from Kazakhstan (7), Turkey (2), Greece (2), Jordan (1), Vietnam (1), Indonesia (1), Romania (1), and multinational survey contexts (5). In addition, five supplementary theoretical works were included to support the theoretical synthesis: the foundational texts by Vygotsky L. S. (1978) [1], Krashen S.D. (1982) [2], and Tomlinson C.A. (2001) [3], the Common European Framework of Reference for Languages (Council of Europe, 2020 [24]), and the readability framework developed by Kincaid J. P. et al. (1975) [25].

This study employs three analytical methods. The first is comparative analysis. All 20 selected sources underwent a full reading and systematic comparison according to the following parameters: research focus, theoretical base, educational context and population, and key findings. Special attention went to the four studies that address AI-based text adaptation as their primary focus, namely Murphy Odo D. (2025) [15], Çelik F. et al. (2024) [16], Aldamen H. et al. (2026) [17], an Nguyen T. Q. Y. (2024) [18], as these underwent side-by-side examination in order to identify points of agreement, divergence, and unresolved questions. Kazakhstani sources additionally underwent comparison with international research to map geographic and contextual gaps in the existing body of knowledge.

The second method is classification. Following the comparative reading, all 20 sources underwent systematic categorisation according to the type of research gap they represent. This process resulted in four gap categories, namely theoretical, empirical, population, and methodological, each reflecting a specific dimension in which the existing literature does not adequately address the use of AI for text adaptation in EFL middle school settings.

The third method is theoretical synthesis, applied in a predominantly deductive manner. The three frameworks most relevant to the research question, namely the Zone of Proximal Development by Vygotsky L.S. (1978), the Input Hypothesis by Krashen S.D. (1982), and the Differentiated Instruction model by Tomlinson C.A. (2001), underwent individual analysis before their subsequent integration into a unified conceptual framework. This approach involved two stages: abstraction, that is, identifying the core principle of each theory most directly applicable to AI-based text adaptation, and deduction, that is, applying these general theoretical principles to the specific context of Grade 7 EFL learners in Kazakhstan.

The choice of a theoretical methodology rests on two primary reasons. First, the field of AI-based text adaptation in EFL remains at an early stage of development, and no unified evaluation framework currently exists; establishing such a framework serves as an essential precondition for any credible empirical investigation. Second, a systematic literature review combined with theoretical synthesis represents a well-established and valued research method in applied linguistics, making the research process transparent, structured, and reproducible. The absence of empirical data admittedly constitutes a limitation of the present study. However, this limitation is intentional and strategically motivated: the proposed theoretical framework explicitly serves as the foundation for the subsequent empirical phase of research.

Results and Discussion

The present section reports and discusses the principal outcomes of the systematic critical literature review and theoretical synthesis conducted in this study. In line with the methodological approach described in Methodology, the results consist of a proposed theoretical evaluation framework derived from three established theoretical models, supported by a comparative analysis of prior studies. The section first presents this framework and the comparative analysis, then discusses how it closes the identified research gaps, and concludes with its practical, ethical, and methodological implications.

The central result of the present study is a unified theoretical framework for evaluating the quality of AI-generated text adaptations in EFL middle school settings. The framework emerges from the systematic integration of three foundational models: the Zone of Proximal Development (ZPD) by Vygotsky L. S. [1], the Input Hypothesis by Krashen S.D. (1982) [2], and the

Differentiated Instruction (DI) framework by Tomlinson C.A. (2001) [3]. Each theory is well established independently in applied linguistics, yet, as the comparative analysis below demonstrates, the literature lacks a single operationalised framework combining them – the primary theoretical contribution of this article.

Each theory contributes a distinct, complementary dimension. The ZPD establishes the pedagogical rationale: an AI-adapted text is appropriate only if its difficulty falls within the learner's zone of proximal development – above the current independent level, yet reachable with support – with AI functioning as a digital scaffold. The Input Hypothesis (i+1) operationalises this in linguistically measurable terms, requiring comprehensible input one level above the learner's current CEFR competence, assessed through Flesch-Kincaid readability scores as a quantifiable proxy for text complexity. The DI model addresses the structural dimension: in a mixed-proficiency classroom a single AI-generated adaptation is pedagogically insufficient, so the framework requires a minimum of three parallel versions (A1, A2, B1) from a single source text, without loss of core meaning.

Together the three criteria form a hierarchically integrated system rather than three independent measures: difficulty alignment sets the pedagogical goal, readability alignment supplies the linguistic instrument for reaching it, and differentiation completeness ensures the full proficiency range of a typical Grade 7 Kazakhstani EFL classroom is addressed. Table 1 summarises the complete framework.

Table 1 – Unified Theoretical Evaluation Framework for AI-Based Text Adaptation in EFL Middle School Settings

Theoretical Source	**Core Principle**	**Application to AI Adaptation**	**Operationalisation**	**Evaluation Criterion**
Vygotsky L. S. (1978) [1] – eff Zone of Proximal Development	Learning is most ective when dig input is slightly above the learner's current independent ability, with appropriate support.	AI acts as a ital pos scaffold. The adapted text must challenge the learner without exceeding their capacity for comprehension.	Text difficulty is itioned one CEFR ali level above the learner's current independent level (e.g., A2 text for an A1 learner).	Difficulty gnment: the text must fall within the learner's ZPD as defined by CEFR level descriptors.
Krashen S. D. (1982) [2] – acq Input Hypothesis (i+1)	Language uisition voc occurs when learners receive comprehensible input marginally above their current competence	AI-generated abulary load, rea sentence structure, and grammatical complexity must correspond to the target CEFR level.	The Flesch-Kincaid (FK) dability score serves ali as a measurable proxy for CEFR-level linguistic complexity.	Readability gnment: the FK score of the adapted text must fall within the range associated with the target CEFR level.
Tomlinson C. A. (2001) [3] – tea Differentiated Instruction	Effective ching in a m mixed-ability classrooms requires materials systematically adjusted to different learner groups' readiness levels.	AI must generate inimum of thr three parallel versions of the same source text, each targeting a distinct proficiency group.	The framework mandates ee versions (A1, A2, com and B1), representing the typical proficiency range of Kazakhstani Grade 7 EFL learners.	Differentiation pleteness: all versions must be internally consistent and preserve the core meaning of the original text.

CEFR = Common European Framework of Reference for Languages (Council of Europe, 2020 [24]); FK = Flesch-Kincaid readability formula (Kincaid et al., 1975 [25]); ZPD = Zone of Proximal Development; DI = Differentiated Instruction.

To validate this synthesis against the empirical record, the study compares the four most relevant prior works across five parameters (Table 2).

Table 2 – Comparative Overview of Empirical Studies on AI-Based EFL Text Adaptation

Study	**Theoretical Base**	**Population**	**AI Tool**	**Key Finding**	**Limitation Addressed by Present Framework**
Murphy Odo D. (2025) [15]	Not stated	L2 readers (mixed levels) so	Automatic simplification ftware hi	Simplified texts benefited gher-proficiency ve readers only; no benefit for lower-proficiency group	Lack of differentiated rsions for multiple levels; no readability evaluation criteria
Çelik F. et al. (2024) [16]	Krashen S. D. (i+1) st	105 Turkish university EFL udents	ChatGPT	Significant gains in comprehension and ferencing; no op reduction in reading anxiety	Input Hypothesis cited but not erationalised through CEFR or readability metrics
Aldamen H. et al. (2026) [17] Th	Krashen S. D. (i+1); Cognitive Load eory	48 Jordanian university EFL students	Generative AI (unspecified)	Lower-proficiency learners benefited significantly; gher-proficiency li showed no disadvantage	Theory referenced but not systematically nked to evaluation criteria; tertiary context only
Nguyen T. Q. Y. (2024)	Not stated	EFL teachers in Vietnam	ChatGPT	ChatGPT shows potential for text adaptation; output lacks consistency; no evaluation framework proposed	No theoretical framework; no replicable criteria for quality assessment

EFL = English as a Foreign Language; CEFR = Common European Framework of Reference for Languages.

As Table 2 shows, only two of the four studies cite an explicit theoretical framework, and both rely solely on the Input Hypothesis without connecting it to CEFR descriptors or readability metrics; none addresses secondary school learners; and none proposes or applies a multi-criteria evaluation framework. These patterns confirm, in concrete terms, the theoretical, empirical, and methodological gaps identified in the literature review.

No prior study in the reviewed literature integrates the ZPD by Vygotsky L. S., the Input Hypothesis by Krashen S. D., and the DI model by Tomlinson C. A. into a unified evaluation framework for AI-based text adaptation. Aldamen H. et al. (2026) [17] and Çelik F. et al. (2024) [16] both invoke the Input Hypothesis as theoretical justification, but neither operationalises it through explicit, measurable criteria – so the claim that AI-adapted texts represent genuine i+1 input remains assumed rather than demonstrated. By providing three operationalised criteria, each linked to a specific theoretical principle and measurement method, the present framework resolves this inconsistency. The same university-level concentration of existing studies also leaves secondary contexts without a criteria-based evaluation instrument incorporating CEFR alignment and readability metrics. The finding by Murphy Odo D. (2025) [15] – that simplified texts benefited higher- but not lower-proficiency readers – is telling here: a single AI-generated version, however simplified, may not match the ZPD of the least proficient learners in a mixed-ability group. The differentiation-completeness criterion, requiring a minimum of three parallel versions at distinct CEFR levels, targets precisely this outcome, offering a theoretical account of one of the field's key unresolved findings.

The population gap is equally clear: existing Kazakhstani studies address teacher attitudes, pre-service readiness, or university-level outcomes almost exclusively. The only study combining differentiated instruction with reading in this context, Abildina S. et al. (2024) [13], concerns primary learners and makes no reference to AI, while Talapova A. K. and Yalkonova Y. D. (2025) [6] involve secondary students aged 15-17 but focus on general AI integration rather than text-adaptation quality.

No prior work situates an evaluation framework within the trilingual, centrally curriculated Kazakhstani secondary EFL context – an absence this article addresses directly. Methodologically, the field also lacks a replicable analytical infrastructure, a limitation explicitly noted by Nguyen T. Q. Y. (2024) [18]; although systematic review combined with theoretical synthesis is well established in applied linguistics generally, it remains absent from AI text-adaptation research specifically. The methodological contribution of this article is therefore twofold: demonstrating the value of a purely theoretical approach at an early stage of a developing field, and producing a structured instrument that future empirical work can apply transparently and reproducibly.

The findings carry direct practical implications for EFL teachers in multilevel classrooms, particularly in Kazakhstan. The three evaluation criteria give teachers a concrete, theory-grounded checklist for assessing AI-generated text adaptations before classroom use, replacing subjective judgement or uncritical reliance on AI output with a principled quality-assurance procedure integrated into lesson preparation – directly relevant given Nogaibayeva A. and Yersultanova G. (2025) [4] and Fazilova A. and Kayip A. (2025) [5], who found that Kazakhstani teachers recognise AI's potential but lack tools to assess it systematically. The framework may also serve as a starting point for teacher training: Sultan Y., Dautova G., and Dalle J. (2025) [12] found that Kazakhstani EFL students possess high AI literacy but only moderate practical application skills, a gap that structured evaluation training could help close.

The integration of AI into classroom material design also raises ethical considerations. AI-generated adaptations may introduce biases in cultural representation, gender, or social context, particularly given large language models' predominantly Western English-language training data – a risk in Kazakhstan's multilingual, post-Soviet educational environment, where simplified texts may still be culturally misaligned with learners' lived experience. Uncritical use also risks reducing teachers' professional agency in favour of automated output, a concern Nogaibayeva A. and Yersultanova G. (2025) [4] documented directly among Kazakhstani teachers. The proposed framework targets precisely this balance, positioning AI adaptation as a tool that supplements rather than replaces teacher expertise.

Several limitations warrant transparency. The framework remains theoretical and requires testing against actual AI-generated text samples; the Flesch-Kincaid formula was calibrated for native-speaker texts rather than EFL/L2 contexts, introducing a degree of approximation; CEFR descriptors do not capture all dimensions of text complexity, such as pragmatic coherence or background-knowledge load; and the review's 20 sources, drawn from three databases, may not fully represent non-English-language scholarship. These limitations point to two directions for future research: applying the three criteria to actual AI-generated adaptations of Grade 7 Kazakhstani EFL reading materials, using Flesch-Kincaid analysis and CEFR-aligned vocabulary profiling; and testing the validated framework in a quasi-experimental classroom study with A1, A2, and B1 learners, using pre- and post-tests of reading comprehension and engagement as key outcomes. These two phases constitute the planned empirical continuation of the present work.

Conclusion

The present article aimed to develop a theoretically grounded, criteria-based framework for evaluating AI-generated text adaptations in EFL middle school settings. To this end, the study carried out a systematic critical literature review of 20 peer-reviewed sources, supported by comparative analysis and theoretical synthesis. The principal result is a unified three-criterion evaluation instrument comprising difficulty alignment, readability alignment, and differentiation completeness, derived from the integration of the Zone of Proximal Development by Vygotsky L. S. (1978), the Input Hypothesis by Krashen S. D. (1982), and the Differentiated Instruction model by Tomlinson C. A. (2001). Comparative analysis of four key empirical studies confirmed that no prior work has integrated all three frameworks into a single operational model, addressed secondary school learners, contextualised findings within the Kazakhstani EFL setting, or proposed a replicable criteria-based evaluation methodology. The present article addresses each of these gaps directly.

In terms of practical implications, the framework offers EFL teachers, particularly those working in multilevel Kazakhstani classrooms, a principled quality assurance procedure for evaluating AI-generated materials before classroom use. Rather than applying AI output uncritically or relying solely on subjective professional judgement, teachers can use the three criteria as a structured checklist integrated into lesson preparation. The framework also holds direct relevance for teacher training programmes, where it can serve as a concrete pedagogical tool for developing AI literacy alongside content-specific evaluation competence. Importantly, the framework is immediately applicable in its current theoretical form: teachers can use the three criteria as a pre-use checklist when evaluating any AI-generated reading material, without waiting for empirical validation of the full instrument.

The most significant limitation of the study is its theoretical nature, as the framework currently lacks validation against actual AI-generated text samples. This gap suggests two clear directions for future research. First, future studies should apply the three criteria to AI-generated adaptations of authentic Grade 7 EFL materials from the Kazakhstani national curriculum, using Flesch-Kincaid analysis and CEFR-aligned vocabulary profiling as primary instruments. Second, the next phase of research will involve testing the framework in a quasi-experimental classroom study with Kazakhstani middle school learners at A1, A2, and B1 proficiency levels. This approach will allow for measuring reading comprehension and learner engagement as primary outcome variables. These two phases constitute the planned empirical continuation of the present work, aiming to produce a practically tested, locally contextualised instrument for AI-assisted differentiated reading instruction.

References

1. Vygotsky L.S. Mind in society: The development of higher psychological processes. – Cambridge: Harvard University Press, 1978. – 159 p.
2. Krashen S.D. Principles and practice in second language acquisition. – Oxford: Pergamon Press, 1982. – 202 p.
3. Tomlinson C.A. How to differentiate instruction in mixed-ability classrooms. – Alexandria: ASCD, 2001. – 117 p.
4. Nogaibayeva A., Yersultanova G. ‘It is not the same as a classroom teacher’: A qualitative study of foreign language teachers’ perspectives on artificial intelligence-supported tools in Kazakhstan // Contemporary Educational Technology. – 2025. – Vol. 17. – № 4. – P. ep599.
5. Fazilova A., Kayip A. EFL teachers’ attitudes towards the use of AI generated lesson plans in Kazakhstan // SDU Bulletin: Pedagogy and Teaching Methods. – 2025. – Vol. 71. – № 2. – P. 30-46.
6. Talapova A.K., Yalkonova Y.D. AI tools in English language teaching: Opportunities and challenges // Research Reviews (Pedagogical Sciences). – 2025. – № 11.
7. Uteshova A. Integrating artificial intelligence tools in English language teaching: Global practices and local adaptation // Eurasian Science Review. – 2025. – Vol. 3. – № 5. – P. 18-28.
8. Mendekenova A., Aukhadieva Z., Zhanikeyeva D. Beyond the hype: AI outperforms VR/AR and gamification in foreign language education // Journal of Information Technology Education: Innovations in Practice. – 2025. – Vol. 24. – P. 029.
9. Laoha R., Chomthong W., Pongpanich W. Artificial intelligence and English as a foreign language (EFL) teachers’ competencies: A systematic review // Higher Education Studies. – 2025. – Vol. 15. – № 3. – P. 262-292.
10. Kemelbekova Z., Degtyareva X., Seidaliyeva G. AI in teaching English as a foreign language: Effectiveness and prospects in Kazakh higher education // XLinguae. – 2024. – Vol. 17. – № 1. – P. 69-83.
11. Mektepbergenov N., Nazanova S., Makhambetov Y., Orynbayeva L., Bayaliyev R., Nurgaliyeva S. Pre-service teachers’ readiness to use Artificial Intelligence: Evidence from Kazakhstan // International Journal of Information and Education Technology. – 2025. – Vol. 15. – № 12. – P. 2700-2706.

12. Sultan Y., Dautova G., Dalle J. Examining the relationship among artificial intelligence literacy, cultural literacy, and intercultural communication proficiency of philology students // Journal of Ethnic and Cultural Studies. – 2025. – Vol. 12. – № 5. – P. 345-362.
13. Abildina S., Sarsekeyeva Z., Mukhametzhanova A., Kopbalina K., Nurgaliyeva S. Enhancing reading literacy among elementary school learners in Kazakhstan: The application and effectiveness of modern teaching techniques // Journal of Infrastructure, Policy and Development. – 2024. – Vol. 8. – № 8. – 5905 p.
14. Azamatova A., Bekeyeva N., Zhaxylikova K., Sarbassova A., Ilyassova N. The effect of using artificial intelligence and digital learning tools based on project-based learning approach in foreign language teaching on students' success and motivation // International Journal of Education in Mathematics, Science and Technology. – 2023. – Vol. 11. – № 6. – P. 1458-1475.
15. Murphy Odo D. The effect of automatic text simplification on L2 readers' text comprehension // Applied Linguistics. – 2023. – Vol. 44. – № 6. – P. 1030-1046.
16. Çelik F., Yangın Ersanlı C., Arslanbay G. Does AI simplification of authentic blog texts improve reading comprehension, inferencing, and anxiety? A one-shot intervention in Turkish EFL context // International Review of Research in Open and Distributed Learning. – 2024. – Vol. 25. – № 3. – P. 287-303.
17. Aldamen H., Alnemrat A., Al-Deaibes M., Alsharefeen R. Using AI to adapt reading input for mixed-proficiency EFL learners // Frontiers in Education. – 2026. – Vol. 11. – P. 1737903.
18. Nguyen T. Q. Y. Unraveling the potential of ChatGPT: Investigating the efficacy of reading text adaptation // Proceedings of the 20th AsiaCALL International Conference (AsiaCALL2023). – 2023. – Vol. 4. – P. 159-169.
19. Rapti C., Panagiotidis P. Teachers' attitudes towards AI integration in foreign language learning: Supporting differentiated instruction and flipped classroom // European Journal of Education. – 2024. – Vol. 7. – № 2. – P. 88-104.
20. Karataş F., Eriçok B., Tanrikulu L. Reshaping curriculum adaptation in the age of artificial intelligence: Mapping teachers' AI-driven curriculum adaptation patterns // British Educational Research Journal. – 2025. – Vol. 51. – № 1. – P. 154-180.
21. Chitez M., Csürös K., Rogobete R. Upgraded literacy: Teacher training approaches to integrating corpus data and AI tools for school text readability adaptation // Applied Corpus Linguistics. – 2025. – Vol. 5. – P. 100181.
22. Rahmawati Y., Fitriati S. W., Wahyuni S. A four-stranded approach to AI-assisted pre-class learning modules: EFL teachers' practices in designing daily conversation content for flipped classrooms // Language Teaching Research Quarterly. – 2025. – Vol. 49. – P. 90-112.
23. Woo D. J., Guo K., Susanto H. Exploring EFL students' prompt engineering in human–AI story writing: An activity theory perspective // Interactive Learning Environments. – 2025. – Vol. 33. – № 1. – P. 863-882.
24. Council of Europe. Common European framework of reference for languages: Learning, teaching, assessment. Companion volume. – Strasbourg: Council of Europe Publishing, 2020. – 272 p.
25. Kincaid J. P., Fishburne R. P., Rogers R. L., Chissom B. S. Derivation of new readability formulas for Navy enlisted personnel [Electronic resource]. – Millington, TN: Navy Research Branch, 1975. – URL: https://apps.dtic.mil/sti/citations/ADA006655 (date of access: 04.2026).

С.И. Ихсанова[1*], Ж.У. Калимова[2]
Шәкәрім университеті,
071412, Қазақстан Республикасы, Семей қ., Глинка к-сі, 20 А
[*]ORCID: 0009-0007-5210-8194
[*]e-mail: sofyaikhsanova578@gmail.com

АҒЫЛШЫН ТІЛІ САБАҚТАРЫНДА МӘТІНДЕРДІ АВТОМАТТЫ САРАЛАУ ҮШІН ЖАСАНДЫ ИНТЕЛЛЕКТ ҚҰРАЛДАРЫН ПАЙДАЛАНУ

Аңдатпа. Ғылыми мақаланың мақсаты – орта мектеп деңгейінде шетел тілі ретінде ағылшын тілін оқытуда жасанды интеллект арқылы бейімделген мәтіндердің сапасын бағалауға арналған теориялық тұрғыдан негізделген, критериялық бағалау шеңберін әзірлеу. Зерттеу теориялық сипатта орындалған және сараптамадан өткен 20 дереккөздің (2023-2026 жж.) жүйелі сыни шолуына, сондай-ақ салыстырмалы талдау мен теориялық синтезге негізделген. Жұмыста үш аналитикалық тәсіл қолданылды: іріктелген дереккөздерді салыстырмалы түрде зерттей отырып оқу, зерттеу олқылықтарын жіктеу және дедуктивті теориялық интеграция әдісі. Зерттеудің негізгі нәтижесі – деңгейлік күрделілікке сәйкестік, оқылым түріне сәйкестік және саралау толықтығы сияқты үш критерийді қамтитын бірыңғай бағалау жүйесінің айқындалғындығы. Атап айтқанда, бірыңғай бағалау жүйесі Л.С. Выготскийдің жақын даму аймағы тұжырымдамасын, С.Д. Крашеннің түсінікті енгізу гипотезасын және К.Э. Томлинсонның саралап оқыту моделін біріктіру негізінде жасалып, CEFR деңгей дескрипторлары мен Флеш-Кинкейд оқылымдылық көрсеткіштерін қолдану арқылы анықталды. Төрт негізгі эмпирикалық зерттеуді салыстырмалы талдау бұл саладағы бірде-бір бұрынғы жұмыстың үш теориялық шеңберді бірге қарастырылмағандығын, орта мектеп оқушыларына бағытталмағанын және қайталануға болатын бағалау әдіснамасының бұдан бұрын ұсынылмағанадығын растайды. Ұсынылған шеңбер әртүрлі деңгейдегі сыныптарда жұмыс істейтін ағылшын тілі мұғалімдеріне жасанды интеллект материалдарын сабақта қолданардан бұрын бағалау қағидасына негізделген түрін ұсынады және зерттеудің алдағы эмпирикалық кезеңіне арналған құрылымдық негіз болып отыр.

Тірек сөздер: жасанды интеллект, мәтінді бейімдеу, саралап оқыту, цифрлық ресурстар, ағылшын тілінде оқуға үйрету.

С.И. Ихсанова[1*], Ж.У. Калимова[2]
Шәкәрім университет,
071412, Республика Казахстан, г. Семей, ул. Глинки, 20 А
[*]ORCID: 0009-0007-5210-8194
[*]e-mail: sofyaikhsanova578@gmail.com

ИСПОЛЬЗОВАНИЕ ИНСТРУМЕНТОВ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА ДЛЯ АВТОМАТИЧЕСКОЙ ДИФФЕРЕНЦИАЦИИ ТЕКСТОВ НА УРОКАХ АНГЛИЙСКОГО ЯЗЫКА

Аннотация. Целью данной статьи является разработка теоретически обоснованной, критериальной рамки для оценки качества ИИ-адаптированных текстов в условиях обучения английскому языку как иностранному на уровне средней школы. Исследование носит теоретический характер и основано на систематическом критическом обзоре 20 рецензируемых источников (2023-2026 гг.), дополненном сравнительным анализом и теоретическим синтезом. В ходе работы применялись три аналитические процедуры: сравнительное чтение отобранных источников, классификация исследовательских пробелов и дедуктивная теоретическая интеграция. Основным результатом исследования является единая система оценки, состоящая из трех критериев: соответствия уровня сложности, соответствия параметров удобочитаемости и полноты дифференциации. Данная система разработана на основе интеграции концепции зоны ближайшего развития Л.С. Выготского, гипотезы входного материала С.Д. Крашена и модели дифференцированного обучения К.Э. Томлинсона, а также операционализирована через дескрипторы уровней CEFR и индексы удобочитаемости Флеша-Кинкейда. Сравнительный анализ четырёх ключевых эмпирических исследований подтвердил, что ни одна из предшествующих работ не объединяла все три теоретические рамки, не была направлена на учащихся средней школы и не предлагала воспроизводимой методологии оценки в данной области. Предложенная структура предоставляет учителям английского языка в разноуровневых классах обоснованную процедуру обеспечения качества для оценки материалов, созданных искусственным интеллектом, перед их использованием на уроках, а также служит структурированной основой для эмпирического этапа исследования.

Ключевые слова: искусственный интеллект, адаптация текста, дифференцированное обучение, цифровые ресурсы, обучение чтению на английском языке

Information about the authors

Zhibek Kalimova – candidate of philological sciences, Associate Professor, Department of Foreign and Russian languages «Shakarim University», Republic of Kazakhstan, Semey, e-mail: zhibek_15_10@mail.ru, ORCID: https://orcid.org/0000-0001-8391-8340.

Sofya Ikhsanova* – master's student, Shakarim University, Republic of Kazakhstan, Semey, e-mail: sofyaikhsanova578@gmail.com, ORCID: https://orcid.org/0009-0007-5210-8194.

Авторлар туралы мәліметтер

Калимова Жібек Усембайқызы – ф.ғ.к., «Шәкәрім университеті» КеАҚ шетел және орыс тілдері кафедрасының доценті, Қазақстан Республикасы, Семей, e-mail: zhibek_15_10@mail.ru, ORCID: https://orcid.org/0000-0001-8391-8340.

Ихсанова Софья Ильдаровна* – Шәкәрім университетінің магистранты, Қазақстан Республикасы, Семей, e-mail: sofyaikhsanova578@gmail.com, ORCID: https://orcid.org/0009-0007-5210-8194.

Сведения об авторах

Калимова Жибек Усембаевна – к.ф.н., доцент кафедры иностранного и русского языков НАО «Шәкәрім университет», Республика Казахстан, Семей, e-mail: zhibek_15_10@mail.ru, ORCID: https://orcid.org/0000-0001-8391-8340.

Ихсанова Софья Ильдаровна* – магистрант Шәкәрім университета, Республика Казахстан, Семей, e-mail: sofyaikhsanova578@gmail.com, ORCID: https://orcid.org/0009-0007-5210-8194.

Received 11.08.2026
Accepted 27.09.2026