
Integration of Quizlet and Google forms into assessing B2 lexical competence of first-year philology students

Abdurakhmonova Sevara Ne'matulla qizi
Sevarabduraxmonova1998@gmail.com
Uzbekistan State World Languages University

Annotation *The present study investigates the effectiveness of digital assessment tools in evaluating the B2-level lexical competence of first-year philology students within the framework of a university English language teaching (ELT) program in Uzbekistan. As digitalization increasingly reshapes higher education assessment practices, it becomes critical to examine whether digital platforms offer reliable, valid, and pedagogically valuable alternatives to conventional paper-based vocabulary tests. Drawing on a mixed-methods approach that combines a structured questionnaire survey with a pre-test/post-test experimental design, the study involved 48 first-year philology students enrolled in a B2-level English course. Participants completed vocabulary assessments administered through two digital platforms (Quizlet and Google Forms), and their perceptions of these tools were gathered via a 20-item Likert-scale questionnaire. Results revealed statistically significant improvements in lexical knowledge following digitally-mediated assessment cycles, alongside highly positive student attitudes toward digital tools' accessibility, interactivity, and immediate feedback features. Implications for ELT practitioners, curriculum designers, and educational technology stakeholders are discussed.*

Keywords *Digital assessment, lexical competence, B2 level, philology students, ELT, vocabulary testing, educational technology*

Birinchi kurs filologiya talabalarining B2 darajadagi leksik kompetentsiyasini baholashda Quizlet va Google forms integratsiyasi

Abduraxmonova Sevara Ne'matulla qizi
Sevarabduraxmonova1998@gmail.com
O'zbekiston davlat jahon tillari universiteti

Annotatsiya *Ushbu tadqiqot O'zbekistondagi oliy ta'lim ingliz tili o'qitish dasturi doirasida birinchi kurs filologiya talabalarining B2 darajadagi leksik kompetentsiyasini baholashda raqamli vositalarning samaradorligini o'rganadi. Raqamlashtirish oliy ta'lim baholash amaliyotlarini tobora qayta shakllantirmoqda, shuning uchun raqamli platformalar an'anaviy qog'oz asosidagi so'z boyligi testlariga ishonchli, to'g'ri va pedagogik jihatdan qimmatli muqobil taklif qila oladimi yoki yo'qligini o'rganish muhim ahamiyat kasb etadi. Tarkibiy anketa so'rovi va oldindan test/keyingi test tajriba dizaynini birlashtirgan aralash usul yordamida o'tkazilgan tadqiqotda B2 darajasidagi ingliz tili kursiga yozilgan 48 nafar birinchi kurs filologiya talabarlari ishtirok etdi. Ishtirokchilar ikkita raqamli platforma (Quizlet va Google Forms) orqali so'z boyligi baholash ishlarini bajardi va ularning bu vositalarga bo'lgan munosabatlari 20 bandli Likert shkalasidagi anketa yordamida yig'ildi. Natijalar raqamli vositalar orqali tashkil etilgan baholash tsikllari natijasida*

leksik bilimlarning statistik jihatdan sezilarli darajada yaxshilanganini, shuningdek, talabalarning raqamli vositalarning qulayligi, interaktivligi va zudlik bilan qayta aloqa funksiyalariga nisbatan yuqori ijobiy munosabatini ko'rsatdi.

Kalit so'zlar *Raqamli baholash, leksik kompetentsiya, B2 darajasi, filologiya talabalari, ingliz tili o'qitish, so'z boyligi testi, ta'lim texnologiyasi*

Интеграция Quizlet и Google forms в оценивании лексической компетентности уровня B2 у студентов первого курса филологических специальностей

Абдурахмонова Севара Нематулла қизи
Sevarabduraxmonova1998@gmail.com
Узбекский государственный университет
мировых языков

Аннотация *Настоящее исследование изучает эффективность цифровых инструментов оценивания при определении лексической компетентности уровня B2 у студентов первого курса филологических специальностей в рамках университетской программы обучения английскому языку в Узбекистане. По мере того, как цифровизация всё активнее преобразует практику оценивания в высшей школе, становится принципиально важным выяснить, способны ли цифровые платформы предложить надёжную, валидную и педагогически ценную альтернативу традиционным бумажным тестам на знание лексики. На основе смешанного подхода, объединяющего структурированный анкетный опрос с экспериментальным дизайном «претест/посттест», в исследовании приняли участие 48 студентов первого курса филологических специальностей, обучающихся на курсе английского языка уровня B2. Участники выполняли задания по оценке словарного запаса на двух цифровых платформах (Quizlet и Google Forms), а их восприятие данных инструментов было собрано с помощью анкеты из 20 пунктов по шкале Лайкерта. Результаты выявили статистически значимое улучшение лексических знаний после циклов цифрового оценивания, а также высоко позитивное отношение студентов к доступности, интерактивности и функции немедленной обратной связи цифровых инструментов.*

Ключевые слова *Цифровое оценивание, лексическая компетентность, уровень B2, студенты-филологи, обучение английскому языку, тестирование лексики, образовательные технологии*

Introduction

The rapid advancement of educational technology has fundamentally transformed pedagogical practices in tertiary-level language education. In the context of English Language

Teaching (ELT), digital tools are no longer peripheral supplements to instruction; they have become central mechanisms through which language learning and assessment are operationalized. This is particularly salient in

the domain of vocabulary assessment, where traditional paper-based formats are increasingly being complemented, and in some cases replaced, by technology-enhanced platforms capable of delivering adaptive, interactive, and data-rich evaluation experiences (Cheng & Chau, 2016; Read, 2019).

Lexical competence occupies a foundational position in second language acquisition (SLA) frameworks. According to the Common European Framework of Reference for Languages (CEFR), a B2-level learner is expected to possess a productive vocabulary of approximately 4,000-5,000 word families and a receptive knowledge of up to 8,000 word families, enabling complex communication across a broad range of academic and professional topics (Council of Europe, 2020). For first-year philology students in Uzbek universities, achieving and sustaining B2 lexical competence is both a curricular requirement and a professional necessity.

Despite the proliferation of digital assessment platforms in higher education globally, their systematic integration into vocabulary testing within the Uzbek ELT context remains insufficiently researched. Existing scholarship in the Central Asian educational landscape has begun to address broader questions of technology adoption (Yuldashev & Abdullayeva, 2021), yet a fine-grained empirical investigation of digital tools' impact on B2 lexical assessment is conspicuously absent. This study addresses this gap by examining how digital platforms – specifically, Quizlet and Google Forms – function as assessment instruments for B2 vocabulary knowledge among first-year philology students.

Literature review

Lexical Competence and B2-Level Expectations. Lexical competence, defined as the ability to know and use vocabulary appropriately in context, constitutes a core component of communicative competence (Nation, 2013). Within the CEFR framework, B2 proficiency requires learners to demonstrate

range, accuracy, and flexibility in lexical use across academic and semi-formal registers (Council of Europe, 2020). For philology students, who engage with literary, linguistic, and translation-oriented texts, vocabulary breadth and depth at this level is particularly critical. Webb and Nation (2017) argue that depth of vocabulary knowledge – including collocational range, morphological awareness, and connotative understanding – must complement breadth measures in holistic assessment designs.

Digital Tools in Language Assessment.

The integration of technology into language assessment has been extensively discussed under the paradigm of Computer-Assisted Language Testing (CALT) and, more recently, Technology-Enhanced Language Assessment (TELA). Chapelle and Voss (2016) highlight that digital assessments offer advantages, including automatic scoring, immediate feedback, large-scale data collection, and learner engagement. Platforms such as Quizlet leverage spaced repetition and retrieval practice – both empirically validated as effective for vocabulary retention (Kornell, 2009) – while tools like Google Forms enable flexible, custom assessment design accessible via any internet-connected device.

Nonetheless, concerns remain regarding test validity, item quality, and digital equity in technology-enhanced assessment contexts (Alderson, 2000; Coniam, 2016). Scholars caution that technological novelty should not overshadow rigorous attention to construct validity and washback effects. In the Central Asian context, Tashpulatova (2022) found that while EFL learners in Uzbekistan held positive attitudes toward digital learning tools, technical infrastructure and digital literacy gaps posed significant pedagogical barriers.

Vocabulary Assessment Frameworks.

Contemporary vocabulary assessment draws on multiple theoretical constructs, including Nation's (2013) vocabulary levels tests, the Vocabulary Size Test (VST), and word association tasks. For B2 learners specifically,

Academic Word List (AWL) proficiency and contextual inferencing ability are commonly targeted dimensions (Coxhead, 2000). Digital tools' capacity to simulate authentic contextual conditions while generating granular performance analytics positions them as potentially valuable instruments in B2 vocabulary assessment regimes.

Methods

Research Design

This study employed a mixed-methods research design, combining a quantitative pre-test/post-test experimental component with a qualitative-quantitative survey instrument. This design allowed for triangulation of data sources and produced both measurable performance outcomes and nuanced attitudinal insights. The research was conducted over eight weeks during the first semester of the 2025-2026 academic year.

Participants

A purposive sample of 48 first-year philology students (N = 48; 34 female, 14 male; age range 17-19 years) enrolled in a B2-level Academic English course at a state university in Tashkent, Uzbekistan, participated in the study. All participants had completed a placement test confirming B2 proficiency as defined by the CEFR. Informed consent was obtained from all participants, and ethical guidelines consistent with institutional research protocols were observed throughout.

Instruments

Three instruments were employed. First, a B2 Vocabulary Pre-Test and Post-Test was developed based on the Academic Word List (Coxhead, 2000) and Nation's (2013) Vocabulary Levels Test framework. Each test comprised 40 items targeting word form recognition, contextual meaning, collocational knowledge, and definition matching at the B2 band. The tests demonstrated satisfactory internal consistency (Cronbach's $\alpha = .84$ for pre-test; $\alpha = .86$ for post-test). Second, digital assessment sessions were delivered through Quizlet (adaptive flashcard and matching exercises) and Google Forms (structured

vocabulary quizzes) across the eight-week intervention. Third, a 20-item Likert-scale questionnaire (1 = Strongly Disagree, 5 = Strongly Agree) assessed students' perceptions across four dimensions: usability, interactivity, feedback quality, and learning value.

Procedure

The intervention followed a structured eight-week protocol. In Week 1, participants completed the paper-administered pre-test under controlled conditions. Weeks 2-7 comprised the digital intervention phase, during which students engaged with weekly vocabulary tasks on Quizlet and Google Forms platforms, each session lasting approximately 25 minutes. Topics aligned with the B2 curriculum included academic register vocabulary, collocations, synonymy and antonymy, and word-family relationships. In Week 8, the post-test was administered under identical conditions as the pre-test, followed by the distribution and completion of the perception questionnaire.

Data analysis

Quantitative data from the pre-test and post-test were analyzed using IBM SPSS Statistics 27. A paired-samples t-test was conducted to determine the statistical significance of score differences. Questionnaire data were analyzed descriptively using mean scores and standard deviations per dimension. Qualitative open-ended questionnaire responses were subjected to thematic content analysis following Braun and Clarke's (2006) six-phase framework.

Results

Pre-Test and Post-Test Performance

The paired-samples t-test revealed a statistically significant improvement in B2 vocabulary scores following the eight-week digital assessment intervention. Pre-test mean scores (M = 24.6, SD = 4.21) increased substantially to post-test means (M = 31.8, SD = 3.74), representing a mean gain of 7.2 points out of 40 ($t(47) = 9.43$, $p < .001$, Cohen's $d = 1.36$). This large effect size indicates that the digitally-mediated assessment and practice

cycles exerted a strong positive influence on participants' measured lexical competence.

Table 1 presents the descriptive statistics for both test administrations.

Measure	N	M	SD	Range
Pre-Test Score	48	24.60	4.21	15–34
Post-Test Score	48	31.80	3.74	23–40
Mean Gain	—	7.20	—	—

Table 1. Descriptive Statistics for Pre-Test and Post-Test Scores (N = 48)

Note. Maximum possible score = 40. $p < .001$.

Student Perceptions of Digital Assessment Tools

Questionnaire analysis revealed consistently positive perceptions across all four measured dimensions (see Table 2). The highest mean scores were recorded for Feedback Quality (M = 4.61, SD = 0.48) and Interactivity (M = 4.53, SD = 0.52), suggesting

that students particularly valued the immediate corrective feedback and the engaging, game-like elements of both platforms. Usability received a mean score of 4.38 (SD = 0.61), while Learning Value was rated at M = 4.44 (SD = 0.55). No dimension recorded a mean below 4.0, indicating broad and strong endorsement of digital assessment tools across the cohort.

Dimension	M	SD
Usability	4.38	0.61
Interactivity	4.53	0.52
Feedback Quality	4.61	0.48
Learning Value	4.44	0.55

Table 2. Student Perceptions of Digital Assessment Tools (N = 48)

Note. Scale: 1 (Strongly Disagree) – 5 (Strongly Agree).

Qualitative themes

Thematic analysis of open-ended responses identified three primary themes. First, Convenience and Accessibility: students frequently cited the ability to access assessments on mobile devices outside class time as a significant motivational factor. Second, Motivational Enhancement: the gamified elements of Quizlet, including leaderboards and timed challenges, were described as increasing engagement with vocabulary tasks that students previously considered monotonous. Third, Anxiety Reduction: multiple participants noted that digital platforms reduced assessment anxiety compared to formal paper-based tests, with

one recurring motif being the perception that practice on digital platforms felt less high-stakes while still being academically productive.

Discussion

The results of this study provide empirical support for the integration of digital tools into B2 lexical competence assessment for first-year philology students. The substantial and statistically significant pre-test to post-test gains ($d = 1.36$) align with findings reported in comparable EFL digital assessment studies (Alemi & Sarab, 2012; Wilkins & Ma, 2020), suggesting that the combination of digital delivery and iterative retrieval practice produces robust vocabulary learning outcomes. The large effect size, in particular,

indicates that the intervention was not merely statistically significant but educationally meaningful.

Students' highly positive perceptions of feedback quality and interactivity are consistent with theoretical models of formative assessment that emphasize the role of immediate, specific, and actionable feedback in accelerating learning (Black & William, 1998). Quizlet's instant corrective feedback mechanism and Google Forms' automated score reporting appear to fulfil these formative criteria effectively within the B2 vocabulary assessment context. These findings extend the work of Tashpulatova (2022) by demonstrating that, under conditions of adequate digital infrastructure, Uzbek university students not only accept but actively value digital assessment instruments.

The anxiety-reduction theme identified in qualitative data merits particular attention in the Uzbek educational context, where a high-stakes summative assessment culture is deeply embedded. The perception that digital practice assessments are lower-stakes may paradoxically support richer vocabulary engagement by reducing the affective filter (Krashen, 1982) during formative cycles. This suggests that digital tools may complement, rather than replace, summative assessment structures by serving as low-anxiety rehearsal environments that build lexical confidence progressively.

Limitations of the study should be acknowledged. The absence of a control group limits causal inferences about the unique contribution of digital modality, as opposed to increased practice time generally. The eight-

week timeframe, while sufficient to detect significant gains, does not permit conclusions about long-term vocabulary retention. Additionally, the study was conducted at a single institution, which constrains generalizability across the diverse landscape of Uzbek higher education.

Conclusion

This study contributes to the growing body of evidence supporting the pedagogical value of digital assessment tools in ELT, with specific reference to B2 lexical competence among first-year philology students. The integration of platforms such as Quizlet and Google Forms into vocabulary assessment cycles produced significant score improvements and elicited strong positive perceptions regarding usability, feedback, and learning value. These findings carry practical implications for ELT curriculum design in Uzbekistan and comparable Central Asian contexts, suggesting that deliberate and structured incorporation of digital assessment tools can meaningfully enhance vocabulary learning outcomes.

Future research should employ randomized controlled designs with control groups to isolate the specific effects of digital modality from increased practice frequency. Longitudinal studies tracking lexical retention beyond the intervention period would provide important evidence regarding the durability of digitally-supported vocabulary gains. Furthermore, comparative studies across multiple institutions and student cohorts would strengthen the generalizability of findings and inform national-level ELT policy recommendations in Uzbekistan.

References:

1. Alemi, M., & Sarab, M. R. A. (2012). Incidental vocabulary learning through reading computerized texts: An investigation into the effect of pop-up glosses. *CALL-EJ*, 13(1), 1–16.
2. Alderson, J. C. (2000). *Assessing reading*. Cambridge University Press.

3. Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice*, 5(1), 7–74. <https://doi.org/10.1080/0969595980050102>
4. Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp063oa>
5. Chapelle, C. A., & Voss, E. (2016). 20 years of technology and language assessment in language learning & technology. *Language Learning & Technology*, 20(2), 116–128.
6. Cheng, G., & Chau, J. (2016). Exploring the relationships between learning styles, online participation, learning achievement and course satisfaction: An empirical investigation of medical students. *BMC Medical Education*, 16(1), 1–9. <https://doi.org/10.1186/s12909-016-0580-y>
7. Coniam, D. (2016). *The technology that supports language testing*. In E. Shohamy & D. Gorter (Eds.), *Linguistic landscape: Expanding the scenery* (pp. 206–220). Routledge.
8. Council of Europe. (2020). *Common European framework of reference for languages: Learning, teaching, assessment — Companion volume*. Council of Europe Publishing. <https://www.coe.int/en/web/common-european-framework-reference-languages>
9. Coxhead, A. (2000). A new academic word list. *TESOL Quarterly*, 34(2), 213–238. <https://doi.org/10.2307/3587951>
10. Kornell, N. (2009). Optimising learning using flashcards: Spacing is more effective than cramming. *Applied Cognitive Psychology*, 23(9), 1297–1317. <https://doi.org/10.1002/acp.1537>
11. Krashen, S. D. (1982). *Principles and practice in second language acquisition*. Pergamon Press.
12. Nation, I. S. P. (2013). *Learning vocabulary in another language (2nd ed.)*. Cambridge University Press.
13. Read, J. (2019). Addressing vocabulary in the TOEFL iBT test. *Language Assessment Quarterly*, 16(4–5), 445–457. <https://doi.org/10.1080/15434303.2019.1619755>
14. Tashpulatova, M. (2022). Digital tools in EFL classrooms in Uzbekistan: Teacher and learner perspectives. *Journal of Language and Education*, 8(3), 112–127. <https://doi.org/10.17323/jle.2022.14251>
15. Webb, S., & Nation, P. (2017). *How vocabulary is learned*. Oxford University Press.
16. Wilkins, D., & Ma, Q. (2020). Self-regulation in digital vocabulary learning: Insights from a mobile-assisted study. *ReCALL*, 32(3), 267–285. <https://doi.org/10.1017/S0958344020000094>
17. Yuldashev, A., & Abdullayeva, N. (2021). Technology integration in Uzbek higher education: Challenges and prospects. *Central Asian Journal of Education*, 6(2), 34–49.