UDC 372.881.111.1:61

USING SPACED REPETITION TECHNOLOGY TO ENHANCE MEDICAL ENGLISH VOCABULARY ACQUISITION AT THE INITIAL STAGE

Zaripbayeva Nasiba Azamat qizi
Urgench State University named after Abu Rayhan Biruni, Urgench, Uzbekistan
E-mail: nasibazaripbaeva@gmail.com
ORCID: https://orcid.org/0009-0001-2744-7027

Abstract. *The article analyses, on the basis of published empirical data, how spaced repetition technology can be used to teach medical English vocabulary to first-year students of medical universities. Sixteen sources were synthesised (two meta-analyses, eleven empirical studies and three corpus studies). Spaced practice outperformed massed practice with a medium-to-large effect (g = 0.58 on immediate posttests). Words relearned at 30-day intervals were recalled after eight years in 15% of cases, compared with 6% for words relearned without intervals. Word pairs that remained in testing were retained at about 80% after one week, against 33–36% when they were dropped from testing. A spaced repetition model for the initial stage is proposed.*

Keywords: *spaced repetition, medical English, lexical competence, Anki, retrieval practice, long-term retention, initial stage.*

BOSHLANG'ICH BOSQICHDA TIBBIY INGLIZ TILI LUG'AT BOYLIGINI O'ZLASHTIRISHNI TAKOMILLASHTIRISHDA INTERVAL TAKRORLASH TEXNOLOGIYASIDAN FOYDALANISH

Annotatsiya. *Maqolada tibbiyot oliy ta'lim muassasalarining 1-kurs talabalariga tibbiy ingliz tili leksikasini o'rgatishda intervalli takrorlash texnologiyasidan foydalanish imkoniyatlari chop etilgan tadqiqotlar ma'lumotlari asosida tahlil qilinadi. Jami 16 ta manba (2 ta meta-tahlil, 11 ta empirik tadqiqot, 3 ta korpus tadqiqoti) umumlashtirildi. Oraliq bilan bajarilgan mashq bir vaqtda to'plab bajarilgan mashqqa nisbatan o'rtacha va yuqori darajadagi samara berishi aniqlandi (darhol o'tkazilgan testlarda g = 0,58). 30 kunlik interval bilan o'zlashtirilgan so'zlar 8 yildan so'ng 15 % hollarda esga olingan, intervalsiz o'rganilganlari esa 6 % hollarda. Muntazam sinovdan o'tkazib borilgan so'z juftliklari bir haftadan keyin taxminan 80 %, sinovdan chiqarilganlari esa 33–36 % darajada saqlangan. Natijalar asosida boshlang'ich bosqich uchun intervalli takrorlash modeli taklif etildi.*

Kalit so'zlar: *intervalli takrorlash, tibbiy ingliz tili, leksik kompetensiya, Anki, faol eslash, uzoq muddatli xotira, boshlang'ich bosqich.*

ИСПОЛЬЗОВАНИЕ ТЕХНОЛОГИИ ИНТЕРВАЛЬНОГО ПОВТОРЕНИЯ ДЛЯ СОВЕРШЕНСТВОВАНИЯ УСВОЕНИЯ МЕДИЦИНСКОЙ АНГЛИЙСКОЙ ЛЕКСИКИ НА НАЧАЛЬНОМ ЭТАПЕ

Аннотация. *В статье на основе опубликованных эмпирических данных анализируются возможности применения технологии интервального повторения при обучении студентов первого курса медицинских вузов английской медицинской лексике. Обобщены 16 источников (2 метаанализа, 11 эмпирических исследований, 3 корпусных исследования). Установлено, что распределённая практика даёт эффект от средней до высокой величины по сравнению с концентрированной (g = 0,58 в непосредственных тестах). Слова, повторявшиеся с интервалом 30 дней, через 8 лет воспроизводились в 15 % случаев, без интервала — в 6 %. Пары слов, которые постоянно проверялись, через неделю сохранялись примерно на 80 %, исключённые из проверки — на 33–36 %. Предложена модель интервального повторения для начального этапа.*

Ключевые слова: *интервальное повторение, медицинский английский язык, лексическая компетенция, Anki, активное припоминание, долговременная память, начальный этап.*

Introduction

For a first-year medical student, the English course begins with words. Before a student can follow a lecture on the circulatory system or read a short case description, he or she has to recognise terms such as artery, haemorrhage, auscultation or benign, together with the general academic words that hold a medical text together. Knowing a word in this sense covers its spoken and written form, its meaning and the way it is used with other words [1]. The volume of such material is large, and most of it is met for the first time during the first two semesters, when students are also learning anatomy, Latin and biology.

The task is not only to learn new words but to keep them. Teachers of English for medical purposes see the same pattern every year: terms prepared for a weekly test are recognised on the day of the test and are lost a month later, so that the final examination becomes a second round of memorisation. Corpus research shows why this loss is costly. The Medical Academic Word List compiled from a corpus of 1,093,011 running words of medical research articles contains 623 word families that account for 12.24% of all tokens in the corpus [2]. A relatively compact set of words therefore carries a large share of the reading load, and forgetting part of it directly reduces comprehension.

The decline of memory over time was described as early as the end of the nineteenth century by H. Ebbinghaus, who also observed that repetitions distributed over time produce stronger memories than repetitions crowded into one sitting [3]. This spacing effect is one of the best-documented findings in the psychology of learning: a quantitative synthesis by N. Cepeda and colleagues located 839 assessments of distributed practice in 317 experiments reported in 184 articles [4]. Spaced repetition technology turns the principle into a working routine. Software such as Anki presents a word card, records whether the learner recalled it and schedules the next review accordingly, so that difficult items return sooner and well-known items return after longer gaps.

The question has direct practical meaning in Uzbekistan. The Resolution of the President of the Republic of Uzbekistan No. PQ-5117 dated 19 May 2021 set the task of raising the quality of foreign language teaching to a new level and of introducing effective methods into all stages of education [5]. In medical universities this means that the English course of the first year has to produce vocabulary that students still hold when they move to clinical disciplines. Spaced repetition tools are free, run on an ordinary smartphone and need no change to the curriculum, which makes them a practical option for this stage of instruction.

The aim of the study is to determine, on the basis of published empirical and corpus data, how effective spaced repetition is for vocabulary retention and to propose a model for applying spaced repetition technology to medical English vocabulary at the initial stage of instruction.

Materials and methods

The study was carried out as a structured literature review guided by the PRISMA 2020 statement. Sources were searched in PubMed, Scopus, Web of Science, ERIC and Google Scholar for the period 1985–2025 with combinations of the terms "spaced repetition", "spacing effect", "distributed practice", "retrieval practice", "vocabulary", "second language", "Anki", "medical students" and "medical word list". Four inclusion criteria were applied: (1) the source is a peer-reviewed article or a meta-analysis; (2) it reports quantitative outcomes of retention, test performance or lexical coverage; (3) the material learned is verbal (foreign-language words, word pairs, facts) or the participants are medical students; (4) the design and sample are described clearly enough to interpret the figures.

The final set contained 16 sources: two meta-analyses [4; 6], eleven experimental, cohort and survey studies [7–17] and three corpus-based word list studies [2; 18; 19]. From each source the design, sample size, learning conditions and main numerical results were extracted exactly as reported by the authors. Percentages and effect sizes (Hedges'

g) were then compared across studies, and the published values were used to build the figures. The comparative method, generalisation and pedagogical modelling were used to derive a spaced repetition model for the first-year medical English course.

Results

Size of the spacing effect. The meta-analysis by S. K. Kim and S. Webb, which brought together 98 effect sizes from 48 experiments with 3,411 second language learners, showed that spacing had a medium-to-large effect on second language learning; on immediate posttests the advantage of spaced over massed practice reached g = 0.58, 95% CI [0.16; 1.00] [6]. Longer spacing was as effective as shorter spacing on immediate tests but more effective on delayed tests, while equal and expanding schedules gave statistically equivalent results [6]. T. Nakata likewise found no reliable advantage of gradually expanding intervals over equal intervals in vocabulary learning [11]. The main studies are summarised in Table 1.

Table 1. Empirical evidence on spaced repetition and vocabulary retention

Study	Design, n	Main finding
Cepeda et al., 2006 [4]	Meta-analysis; 839 assessments, 317 experiments	Distributed practice improves verbal recall; the optimal interval grows with the retention interval
Kim, Webb, 2022 [6]	Meta-analysis; 48 experiments, N = 3,411	Spaced > massed, g = 0.58 (immediate); longer gaps better on delayed tests
Cepeda et al., 2008 [7]	Experiment; >1,350 participants, test delays up to 1 year	Optimal gap ≈ 20–40% of a 1-week delay and 5–10% of a 1-year delay
Bahrick, Phelps, 1987 [8]	Longitudinal; 35 learners, 50 word pairs	Recall after 8 years: 15% (30-day gaps) vs 6% (no gap)
Bahrick et al., 1993 [9]	9-year study; 4 learners, 300 word pairs	13 sessions at 56-day gaps ≈ 26 sessions at 14-day gaps
Karpicke, Roediger, 2008 [10]	Experiment; 40 learners, 40 word pairs	Repeated testing ≈ 80% recall after 1 week; dropping tested items 33–36%
Kim, Webb, 2023 [12]	Experiment; 150 learners, 48 words	Spaced > massed on a 2-week delayed test for both flashcards and fill-in-the-blanks
Deng et al., 2015 [13]	Survey and exam data; 72 medical students	Unique Anki cards predicted USMLE Step 1 score ($B = 5.9\times10^{-4}$, p = 0.024)

Study	Design, n	Main finding
Cooper et al., 2023 [16]	Retrospective cohort; 35 Anki users, 268 non-users	Step 1 failure 2.8% (Anki) vs 10.94% (non-Anki), p > 0.05

Source: compiled by the author from [4; 6–13; 16].

Long-term retention. Short laboratory studies cannot show whether spacing matters for knowledge that a doctor needs for years. The longitudinal data obtained by H. Bahrick and colleagues fill this gap. Thirty-five people who had learned 50 English–Spanish word pairs were tested eight years later; the interval between relearning sessions was the decisive variable (Figure 1). Recall was 15% after training with 30-day intervals, 8% with 1-day intervals and 6% when all sessions took place on the same day; recognition scores under the same conditions were 83%, 80% and 71%, respectively [8; 9].

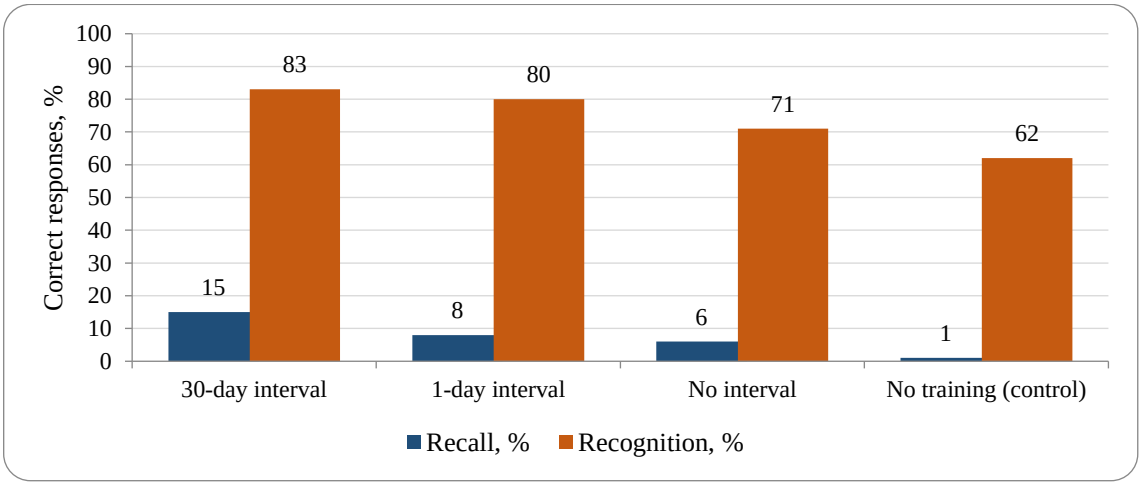


Figure 1. Retention of foreign vocabulary after 8 years by interval between relearning sessions

Source: built by the author based on [8; 9].

A second study in the same programme followed 300 word pairs over nine years and showed that 13 relearning sessions spaced 56 days apart gave retention comparable to 26 sessions spaced 14 days apart [9]. Wider gaps, in other words, allowed the same result with half the number of sessions. The large-scale experiment of Cepeda et al. specified the relation between gap and test delay: when the test followed after one week the best gap was about 20–40% of that delay, and when the test followed after one year it fell to about 5–10% [7].

Retrieval as the working mechanism. Spacing gives its full benefit when each review requires the learner to recall the word, not only to look at it. In the experiment of J. Karpicke and H. Roediger, students learned 40 Swahili–English word pairs under four conditions. When pairs stayed in repeated testing, about 80% were recalled one week

later; when a pair was removed from testing after the first correct answer, recall fell to 36% and 33% (Figure 2) [10]. Repeated study of a pair that had already been recalled added no measurable gain, while repeated testing did.

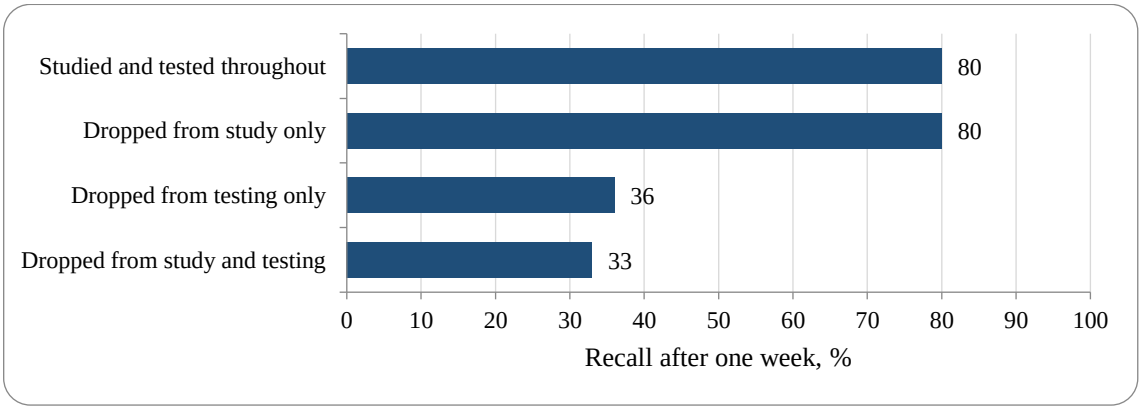


Figure 2. One-week recall of word pairs under four learning conditions

Source: built by the author based on [10]; values are approximate as reported by the authors.

Spaced repetition software in medical education. Studies conducted with medical students point in the same direction, although they are observational. F. Deng and colleagues surveyed 72 students and found that 31% used Anki; after control for MCAT score, grades and psychological factors, the number of unique Anki cards seen remained an independent predictor of the licensing examination score ($B = 5.9\times10^{-4}$, $p = 0.024$) [13]. In a cohort study by M. Gilbert et al., 130 first-year students received training in Anki use; the 78 students who used the application outperformed the 52 non-users on all four examinations analysed, by 6.2% to 12.9% [14]. J. Wothe and co-authors surveyed 165 students, 92 of whom (56%) used Anki daily: daily use correlated with higher Step 1 scores ($p = 0.039$) but not with Step 2 scores, and Anki use was also associated with better sleep quality ($p = 0.01$) [15]. In the study of S. Cooper et al., 2.8% of the Anki cohort ($n = 35$) failed Step 1 against 10.94% of the comparison cohort ($n = 268$) (Figure 3), although the difference did not reach statistical significance and the Anki cohort had a lower grade point average (83.82 vs 85.60) [16].

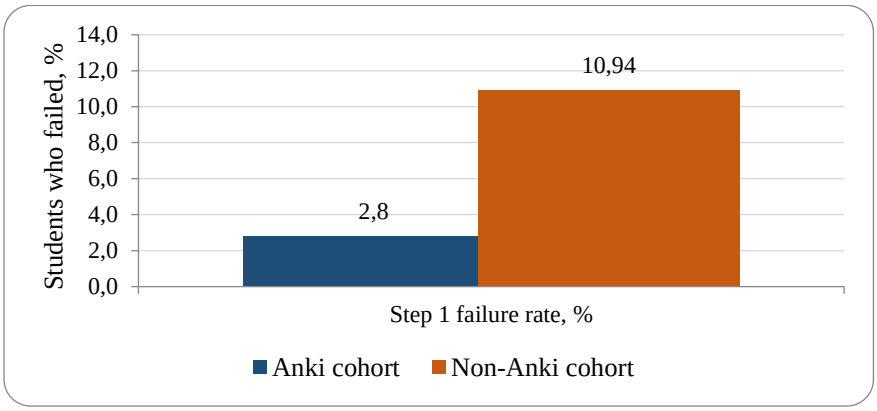


Figure 3. Board examination failure rate in students who did and did not use spaced repetition software

Source: built by the author based on [16].

Adherence is the weak point of the technology. When A. Seibert Hanson and C. Brown made Anki part of the homework grade in a beginning Spanish course, only 14 of 62 students used it at least three days a week, although the number of days of use predicted final performance after control for baseline ability and motivation (b = 0.24, p = 0.022) [17]. The tool, therefore, works only when its use is built into the course and checked by the teacher.

What should be placed in the deck. The selection of words is the second condition of success. Table 2 shows that general academic vocabulary already covers about one tenth of medical research texts, and a discipline-specific list gives slightly higher coverage with a comparable number of word families [2; 18; 19].

Table 2. Lexical coverage of academic word lists in medical texts

Word list	Corpus	Size of list	Coverage
Academic Word List (Coxhead, 2000) [18]	3.5 million words, academic texts of four faculties	570 word families	≈ 10% of tokens
AWL in medical research articles (Chen, Ge, 2007) [19]	190,425 words, medical research articles	292 of 570 families (51.2%) widely present	10.07% of tokens
Medical Academic Word List (Wang et al., 2008) [2]	1,093,011 words, medical research articles of 32 subject areas	623 word families	12.24% of tokens

Source: compiled by the author from [2; 18; 19].

Discussion

The data reviewed show a consistent picture. The benefit of spaced repetition is moderate on the day of learning and grows as the delay increases [6; 8]. This is exactly the condition of a medical student, for whom the words of the first semester must be available at the end of the year and later in clinical work. Two further results matter for course design. First, the precise shape of the schedule is less important than the fact of spacing itself: expanding and equal intervals perform similarly [6; 11]. Second, each review should be an act of recall [10], which is what flashcard software is built for.

For a semester of about 16 weeks the figures of Cepeda et al. [7] suggest that the gap between reviews should gradually reach one to three weeks; this estimate is our interpolation between the values the authors reported for one-week and one-year test delays. Spaced repetition software reaches such gaps automatically if students review daily, because the algorithm lengthens the interval after each successful recall. On this basis a model for the first-year medical English course was developed (Table 3).

Table 3. Model of spaced repetition use in the first-year medical English course

Stage	Content of work	Timing	Form of control
1. Selection	Teacher builds a shared deck by syllabus topics: body systems, symptoms, procedures; MAWL items; Greek-Latin word elements	Before the semester	Deck reviewed by the department
2. Introduction	New words are presented in class in context, with pronunciation and collocations, then added to the deck	Each lesson, 10–15 new cards	Short oral check
3. Independent review	Daily review in Anki or a similar application; first reviews after 1 and 3 days, then by algorithm	15–20 minutes a day	Application statistics
4. Classroom retrieval	Cumulative quiz that includes words of earlier units	Weekly	Written quiz
5. Delayed assessment	Test of words learned 4–8 weeks earlier	Mid-term and final	Delayed vocabulary test

Source: developed by the author on the basis of [6; 7; 10; 17].

Figure 4 shows how the review intervals of a single card expand during the semester under the proposed model. After the lesson in which a word is introduced, it returns on days 1, 3, 7, 14, 30 and 60; a card that the student fails is returned to the short interval.

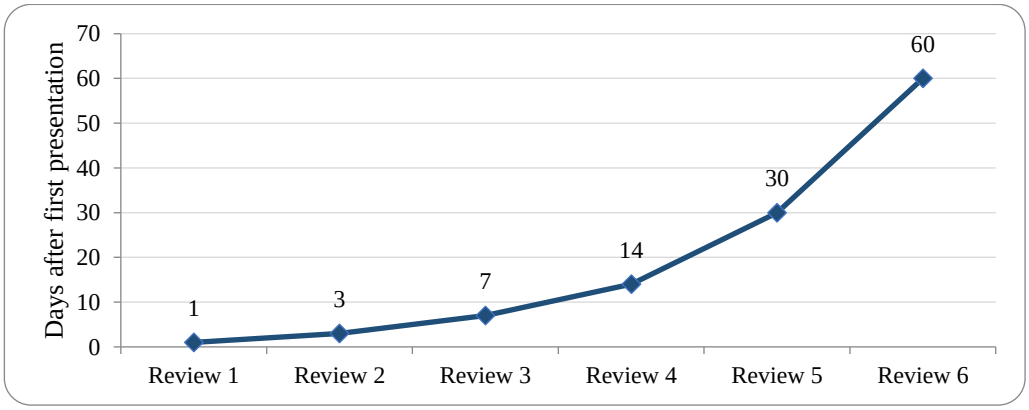


Figure 4. Recommended review schedule for one word card in a 16-week semester

Source: author's model based on [6; 7; 11].

Several limitations must be stated. Most experimental evidence comes from paired-associate learning, in which a learner links a word with its translation; knowledge of collocations and use in speech develops more slowly and needs classroom work in addition to cards. The studies of medical students are correlational, and students who choose spaced repetition may also differ in motivation and study time [13–16]. Finally, low adherence can cancel the advantage of the method [17]. For this reason the model links independent review with weekly classroom retrieval and with delayed tests that count towards the final grade, so that regular review becomes part of the course rather than an optional habit.

Conclusion

1. Published data confirm that spaced practice improves vocabulary learning with a medium-to-large effect (g = 0.58 on immediate tests), and the benefit is largest on delayed tests: words relearned at 30-day intervals were recalled after eight years in 15% of cases against 6% without intervals.

2. Retrieval is the working mechanism of the method: word pairs that remained in testing were retained at about 80% after one week against 33–36% when they were dropped from testing; spaced repetition software combines retrieval and spacing in one daily routine, and in most of the reviewed studies its use by medical students was associated with better examination results.

3. For first-year medical students a five-stage model is proposed: a shared deck built on the syllabus and the Medical Academic Word List, introduction of 10–15 words per

lesson, 15–20 minutes of daily review, a weekly cumulative quiz and delayed vocabulary tests at mid-term and at the end of the semester.

References

1. Nation I.S.P. Learning Vocabulary in Another Language. 3rd ed. – Cambridge: Cambridge University Press, 2022.
2. Wang J., Liang S., Ge G. Establishment of a Medical Academic Word List // English for Specific Purposes. – 2008. – Vol. 27, No. 4. – P. 442–458.
3. Ebbinghaus H. Memory: A Contribution to Experimental Psychology / transl. H.A. Ruger, C.E. Bussenius. – New York: Teachers College, Columbia University, 1913.
4. Cepeda N.J., Pashler H., Vul E., Wixted J.T., Rohrer D. Distributed practice in verbal recall tasks: a review and quantitative synthesis // Psychological Bulletin. – 2006. – Vol. 132, No. 3. – P. 354–380. DOI: 10.1037/0033-2909.132.3.354
5. O‘zbekiston Respublikasi Prezidentining 2021-yil 19-maydagi “O‘zbekiston Respublikasida xorijiy tillarni o‘rganishni ommalashtirish faoliyatini sifat jihatidan yangi bosqichga olib chiqish chora-tadbirlari to‘g‘risida”gi PQ-5117-son qarori. URL: https://lex.uz/docs/5426736
6. Kim S.K., Webb S. The effects of spaced practice on second language learning: a meta-analysis // Language Learning. – 2022. – Vol. 72, No. 1. – P. 269–319. DOI: 10.1111/lang.12479
7. Cepeda N.J., Vul E., Rohrer D., Wixted J.T., Pashler H. Spacing effects in learning: a temporal ridgeline of optimal retention // Psychological Science. – 2008. – Vol. 19, No. 11. – P. 1095–1102. DOI: 10.1111/j.1467-9280.2008.02209.x
8. Bahrick H.P., Phelps E. Retention of Spanish vocabulary over 8 years // Journal of Experimental Psychology: Learning, Memory, and Cognition. – 1987. – Vol. 13, No. 2. – P. 344–349.
9. Bahrick H.P., Bahrick L.E., Bahrick A.S., Bahrick P.E. Maintenance of foreign language vocabulary and the spacing effect // Psychological Science. – 1993. – Vol. 4, No. 5. – P. 316–321. DOI: 10.1111/j.1467-9280.1993.tb00571.x
10. Karpicke J.D., Roediger H.L. The critical importance of retrieval for learning // Science. – 2008. – Vol. 319, No. 5865. – P. 966–968. DOI: 10.1126/science.1152408
11. Nakata T. Effects of expanding and equal spacing on second language vocabulary learning: does gradually increasing spacing increase vocabulary learning? // Studies in Second Language Acquisition. – 2015. – Vol. 37, No. 4. – P. 677–711. DOI: 10.1017/S0272263114000825

12. Kim S.K., Webb S. Does spaced practice have the same effects on different second language vocabulary learning activities? Fill-in-the-blanks versus flashcards // The Modern Language Journal. – 2023. – Vol. 107, No. 4. – P. 944–964.
13. Deng F., Gluckstein J.A., Larsen D.P. Student-directed retrieval practice is a predictor of medical licensing examination performance // Perspectives on Medical Education. – 2015. – Vol. 4, No. 6. – P. 308–313. DOI: 10.1007/s40037-015-0220-x
14. Gilbert M.M., Frommeyer T.C., Brittain G.V., et al. A cohort study assessing the impact of Anki as a spaced repetition tool on academic performance in medical school // Medical Science Educator. – 2023. – Vol. 33, No. 4. – P. 955–962. DOI: 10.1007/s40670-023-01826-8
15. Wothe J.K., Wanberg L.J., Hohle R.D., et al. Academic and wellness outcomes associated with use of Anki spaced repetition software in medical school // Journal of Medical Education and Curricular Development. – 2023. – Vol. 10. DOI: 10.1177/23821205231173289
16. Cooper S., Twardowski N., Vogel M., Perling D., Ryznar R. The effect of spaced repetition learning through Anki on medical board exam performance // International Journal of Medical Students. – 2023. – Vol. 11, No. 4. – P. 271–275.
17. Seibert Hanson A.E., Brown C.M. Enhancing L2 learning through a mobile assisted spaced-repetition tool: an effective but bitter pill? // Computer Assisted Language Learning. – 2020. – Vol. 33, No. 1–2. – P. 133–155. DOI: 10.1080/09588221.2018.1552975
18. Coxhead A. A new academic word list // TESOL Quarterly. – 2000. – Vol. 34, No. 2. – P. 213–238.
19. Chen Q., Ge G. A corpus-based lexical study on frequency and distribution of Coxhead's AWL word families in medical research articles // English for Specific Purposes. – 2007. – Vol. 26, No. 4. – P. 502–514.