

【原著】

Synonym Acquisition and Retention: AI-Assisted Learner-Selected vs Teacher-Provided Synonyms

Craig Anthony Nevitt

同義語の獲得と保持
—AI支援による学習者選択同義語と教師提供同義語の比較—

クレイグ アンソニー ネヴィット

Abstract

This study compares the effectiveness of Teacher-Provided Synonym Sets (TPSS) and AI-Assisted Learner-Selected Synonym Sets (AALSS) in enhancing vocabulary retention among a consistent group of low-proficiency students over a single semester. I administered pre- and post-tests, as well as an anonymous student survey, to collect data. Student perception found through the survey and test performance paint a different picture. Survey results revealed that AI-assisted synonym creation was well-received, with students finding it engaging, motivating, and beneficial for vocabulary learning. The process promoted learner autonomy and collaboration while encouraging active engagement with the material. On a 6-point Likert scale, students rated their engagement positively, with average scores ranging from 4.3 to 4.8 in categories such as interest, motivation, and collaboration. Most students (62%) preferred creating their own synonym sets, citing improved memorization and greater control over their learning. Despite positive student perceptions, test results showed students performed better with TPSS on chapter tests. The final analysis of a delayed post-test and survey finds that the teacher-selected synonyms resulted in a better outcome. The larger improvement in mean scores suggests that the teacher's expertise in selecting appropriate synonyms may have contributed to better student learning. However, both methods led to significant improvements, and the difference in p-values is not substantial. Thus, there appear to be potential benefits of both teacher-selected and student-selected synonyms.

Introduction

This research compares the effectiveness of two variants of a memory vocabulary learning strategy using synonyms. I compare the perceptions and outcomes of Teacher-Provided Synonym Sets (TPSS) and AI-Assisted Learner-Selected Synonym Sets (AALSS) with the same group of low-proficiency students over a single semester regarding vocabulary acquisition and retention.

There are several major categories of vocabulary learning strategies, each of which helps learners acquire new vocabulary in different ways. The VLS I have been focusing on is the memory strategy of using synonyms. I began my research by embedding target synonyms into reading texts, a process known as 'seeding' with synonym sets to improve lower-proficiency student vocabulary acquisition and retention. Through experience and the research of others, I found that the amount of seeding did not produce significant improvement. Boers refers to research by Webb that in a 35-minute reading task, it was necessary to provide 15 encounters with a word for students to recall a verb phrase 56% correctly. Whereas only 10 encounters produced 28% recall. And Boers found that on a test "that exposed them to 5 instances of the same phrase, the students got a score of just 11%" (Boers, 2021, p. 50). Additional vocabulary activities in class and as part of daily practice regimes homework to give students more exposure to the synonym sets failed to engage student or improve their score consistently.

If seeding takes too much effort, often resulting in unnatural occurrences of target language with moderate results, and students are unwilling to engage in daily practice regimes, a different approach is needed. Recent advancements in AI offer promising solutions, particularly for lower-proficiency learners. AI-powered language models, such as chatbots, can generate contextually appropriate synonyms, thus reducing the cognitive load on students and allowing them to engage more actively in the learning process.

In 2023, Nevitt stated that "Matching the lexical profile to lower-proficiency students requires time to create materials and teach students to understand context" (p. 3). The thesaurus, paper and web-based, gives too many synonym choices lacking context. However, recent advancements in AI have opened new possibilities for helping students learn synonyms more effectively. AI-powered language models can understand context and generate relevant synonyms based on the input provided. By leveraging these capabilities, instructors can now create AI-based tools that allow students to find suitable synonyms for target words without being overwhelmed by extensive lists or complex corpus data. For instance, an AI-powered chatbot can be designed to provide students with a set of carefully curated synonyms based on their proficiency level and the context of the target word. This approach simplifies finding appropriate synonyms and enables students to take a more active role in their learning process. By incorporating AI into the synonym learning process, instructors can save time on material creation and focus more on guiding students towards effective vocabulary acquisition strategies. I think AI makes it possible for students to create their own synonym sets, which may lead to improved engagement. This could result in improved vocabulary acquisition and retention. This study aims to investigate the impact of AI-assisted learner-selected synonym sets on vocabulary retention and student confidence. Specifically, my research addresses the following questions:

RQ 1: To what extent does the AI-assisted learner-selected synonym set approach affect the retention of synonyms among low-proficiency EFL students compared to teacher-provided synonym sets?

RQ 2: How does the AI-assisted learner-selected synonym set approach influence the confidence levels of low-proficiency EFL students in retaining synonyms compared to the teacher-

provided synonym sets?

I analyzed the students' perception of the AI-assisted learner-selected synonym set approach through a questionnaire.

RQ 3: What are the perceptions of low-proficiency EFL students regarding the effectiveness and value of using AI to generate their own synonym sets for vocabulary acquisition compared to using teacher-provided synonym sets?

The results suggest that targeted vocabulary instruction, whether through teacher-selected or student-selected synonyms, can lead to significant gains in learners' knowledge and retention of those words. The teacher-selected synonym group had a larger improvement in mean scores compared with the student-selected synonym group. This suggests that the teacher-selected synonyms may have been more effective in improving students' performance.

Synonym Research

For some time, research on vocabulary knowledge did not clearly define what it means to "understand a word." However, I have noticed a positive trend of recent vocabulary researchers focusing on two dichotomies—size or depth of vocabulary knowledge, or receptive or productive knowledge. Charkova and Charkova point out that "The different classifications are not mutually exclusive, but closely linked" (2018, p. 236). Thus, if we are diligent in specifying what we are investigating, researchers can contribute to a more nuanced and comprehensive understanding of the complex processes involved in vocabulary acquisition.

My current line of research is on receptive knowledge, with a focus on synonyms. The concept of receptive knowledge hints at a lack of nuanced understanding between synonyms. I believe, however, that a general understanding of synonyms is a stepping stone toward fluency. I agree that linguistic fluency lies in distinctions. The more complex the reading materials, the more we need to understand distinctions. The ability to understand nuanced texts is an end goal for students of language. However, they need a place to start, and the nuance of distinction is only enjoyable after reaching a modicum of linguist proficiency.

Beginning and intermediate students do not need to understand the distinctions and nuance between words such as work and labor. For a low-proficiency student, understanding work and labor as synonyms is enough information to answer a listening question on TOEIC Part 3:

(M-Au): Did you hear about the new contract our company won for the construction project downtown?

(W-Br): Yes, I did. It's a huge project, but I'm worried about the labor shortage in the area.

(M-Au): That's true, but we're planning to bring in some specialists from out of town to manage the more complex tasks.

(W-Br): That sounds like a good plan. We need to make sure the work is done on time and to a high standard.

What concern does the woman express?

- A) The quality of the necessary materials
- B) The timeline of the project because it is big
- C) The availability of people to do to the work
- D) The high expectations for the project

There are those who would argue that there is more than just the nuance of *work* and *labor* either at the Arendtian distinction of “activity which corresponds to the biological process of the human body” and “activity which corresponds to the unnaturalness of human existence” (Arendt, 2018, p. 7); or at the learners dictionary level of the verb form *work* being everything from “to have a job” to “to do something that involves physical or mental effort” and equivalent verb meanings of *labor*, which range from “to do work” to “to move or proceed with effort” (Britannica Dictionary, n.d.). And, even on this dictionary site for language learners, we naturally find noncorresponding definitions of these terms. Such as a woman in labor, which is an entirely different meaning from a politician working the room. And we are still avoiding the case of phrasal verbs, which is yet another issue to work around. Although words can be synonyms in some cases and not in others, as the test question above illustrates, students do not need to understand every nuance of a word to answer a TOEIC question. The test makers will not use the case of a woman giving birth as a trick to fool test takers. Knowing that they can use the two words interchangeably at times, while also knowing that other meanings exist, is enough to answer the question. Once a student passes the TOEIC with over 900 points, they can choose to split the hairs of work and labor.

Some research does not indicate student proficiency. Thus, it is difficult to find where these researchers place synonyms in their learning order. However, they find students prefer and often enjoy learning synonyms. In their study on Data-driven learning, Asik et al. neither mention their student proficiency, nor analyze when to introduce synonyms, but for their students, “tasks about synonyms and collocations were the mostly favoured” (2016, p. 91). So, based on their understanding of collocations, I assume the students were higher proficiency. One student found creating word webs to be effective and said, “I enjoyed finding the synonyms of the words I have never seen before” (2016, p. 91). Several of my students who have embraced synonym study state that synonyms make it easier to understand the meaning when another word in the set is familiar.

In the Ph.D. thesis by Alharbi on vocabulary acquisition, he creates a vocabulary learning strategy wherein, for “the first step of the sequential VLS processes, participants were asked to first check and connect the new word with as many synonyms as they can. In other words, they relate and make a connection of the new (unknown) words with the known ones” (Alharbi, 2019, p. 67). His students had a higher proficiency, but he instructed his learners to first think of as many synonyms as possible that they already knew. This step in vocabulary learning helps create a mental link between the new word and existing knowledge, making it easier to understand and remember the new vocabulary item. Later in his dissertation, Alharbi puts synonyms as the second step after guessing the meaning from context (Alharbi, 2019, p. 86). In appendix D of the paper, however, his formula returns synonyms to the first action step (Alharbi,

2019, p. 152). This is also the same order Alharbi uses in an earlier paper, but he found that “the group with high language proficiency agreed more on the items selected than those with low language proficiency” (Alharbi, 2015, p. 509). Both groups kept synonyms in the first step and Alharbi advocates for teaching the vocabulary learning strategy “from the early stages of language learning and during vocabulary learning practices” (Alharbi, 2015, p. 509).

Alharbi had student use various vocabulary learning strategies. Similarly, Charkova and Charkova write on vocabulary learning strategies and vocabulary knowledge. They found “that time given to learning words’ multiple meanings and their appropriate synonyms can be very beneficial for expanding learners’ vocabulary knowledge at levels deeper than the surface one” (2018, p. 246). However, they had divided their students into a “limited knowledge” group and a “superior knowledge” group. Whereby they found that the superior knowledge students “reported paying regular attention to target words’ synonyms and antonyms” (2018, p. 245). Nevertheless, when analyzing the data given on page 242 of the paper, I noticed that students in both groups performed somewhat better when studying synonyms compared to collocations. And researchers such as Webb find “that synonymy may facilitate vocabulary learning. The scores were significantly higher for words with known synonyms than those without” (Webb, 2007, p. 130). Furthermore, research has shown that the ability to recognize and use synonyms is a key indicator of vocabulary depth (Read, 2004). Thus, I continue to work to find better ways to teach synonyms and let students learn collocations through incidental acquisition based on the difficulties of covering all collocations. Indeed, some researchers term collocation teaching impossible. Schmitt states, “while it might be possible to explicitly teach a few collocations for a word, good intuitions about all of the words which do and do not collocate with that word would be impossible to teach. Such intuitions could only be acquired by large amounts of exposure” (2010, p. 33). Thus, collocations are more geared toward high-proficiency students. Whereas all learners can benefit from learning synonyms because they foster a broader vocabulary with connections to other words.

AI in vocabulary acquisition and retention

AI is ever evolving. Those who work in the field often say, “today’s AI is the dumbest it’s ever going to be. It’s only going to get smarter and better” (Schlee, 2024, 8:35). Keeping up with the latest AI research is necessary, but is much like drinking from a firehose, only to have it turned off so that you go thirsty while trying to find the next firehose. I have not found papers at the intersection of AI and synonyms for language acquisition. This does not mean that such a paper does not exist, especially considering findings by Compton and Burke on artificial intelligence in higher education, that “language learning was the most common subject domain. This included writing, reading, and vocabulary acquisition” (2023, p. 19). Thus, the search goes on, and through this study, I aim to contribute to research on the use of AI in vocabulary acquisition for language learning. By exploring how AI-assisted learner-selected synonym sets affect vocabulary retention and learner confidence, I hope to provide evidence-based insights that can inform future educational practices and technological developments in language education.

Validation of synonyms on TOEIC tests

To date, I have only relied on my anecdotal notice of synonyms on the TOEIC test. To confirm my suspicions, I reviewed a 2024 TOEIC test sample for the number of times that synonyms were necessary to answer questions. The analysis revealed that many of the question-and-answer word pairs are indeed synonyms, confirming that knowing synonyms is crucial for identifying correct answers on the TOEIC. This finding reinforces the idea that building a strong vocabulary with an emphasis on synonyms can be very beneficial for test-takers.

I reviewed a 2024 sample TOEIC test (Educational Testing Service, 2024) and found compelling evidence for the importance of receptive synonym knowledge in achieving a high score. Out of the 200 questions on the test, 30 (15%) require the test-taker to identify synonyms. This is a significant portion of the test and underscores the need for students to focus on learning synonyms as part of their test preparation strategy.

The review also reveals that synonyms are tested in both the listening and reading sections of the TOEIC. In the listening section, synonyms appear in Part 2's short conversations and Part 3's longer conversations. This means that students must be attentive and listen for words that are repeated or rephrased using synonyms. In the reading section, I found most synonyms in Part 7's long readings. Here, students must carefully read the answer choices and look for synonyms that match the context of the question. Many students state this to be the most difficult section.

Notably, the TOEIC test includes both basic synonyms, such as "happy" and "glad," as well as more advanced ones, like "prominent" and "famous." This suggests that prioritizing learning common, easy synonyms will benefit students as much as trying to learn more sophisticated synonyms, which should, naturally, be included as a student progresses. These findings underscore the vital role of receptive synonym knowledge in achieving success on the TOEIC test. By focusing on learning synonyms in both listening and reading contexts, and across a range of difficulty levels and topics, students can significantly improve their chances of identifying correct answers and boosting their overall test scores.

Methods

This research aims to compare the effectiveness of Teacher-Provided Synonym Sets (TPSS) and AI-Assisted Learner-Selected Synonym Sets (AALSS). I conducted the study with the same group of students in a single semester to ensure consistency in the subjects' abilities and to minimize the impact of potentially confounding variables.

I planned the study to comprise two 5-week reading units, with Chapter 2 emphasizing TPSS and Chapter 3 focusing on AALSS. Additionally, I continued with the same style of daily practice regimes (DPR) though the game-based learning platform, Kahoot! for additional exposure to the words in the synonym sets. I am still trying to improve both efficiency and student engagement in the DPR.

Student test scores and their attempts on the DPR show statistically significant differences ($p < 0.05$). Thus, they appear effective. Unfortunately, 30% of the students do not attempt the DPR,

and only 55%. This is even when attempts affected 10% of their grade.

However, upon further analysis, I found that mean test scores are much higher than the mean DPR attempts, suggesting that, while DPR may have some effect, it is not the primary driver of test performance. The Pearson Correlation between test scores and attempts on the DPR also reveals a weak correlation of 0.35.

The original 5-week reading units plan was disrupted by meteorological and logistical challenges and the need to teach students how to use AI for selecting synonyms. Thus, I had to shorten the AALSS unit to 3 weeks. I administered pre-tests for both chapters. I ensured an even distribution of words from the chapter and controlled the word selection in the Chapter 2 pre-test, while I selected the words from Chapter 3 based on my predictions of student choices.

In Chapter 3, students worked in groups to select words from a three-page text and used AI (ChatGPT, version 4o) to create synonym sets. The class then collaborated to agree on 10 synonym sets to study. I provided students with the following prompt:

Based on the sentence, “ENTER SENTENCE WITH TARGET WORD,” please create a list of synonyms in CEFR level order for the word “TARGET WORD.”

I gave additional guidance to students who needed help with the AI tool, emphasizing the importance of flexibility and persistence when working with AI.

The word selection process in Chapter 3 proved to be chaotic, with one word being chosen that was not present in the text, highlighting the need for careful oversight by the researcher. Students had only 3 weeks to study the AI-generated synonym sets.

Three weeks sufficed to cover the reading materials with additional in class vocabulary activities because I could only create 3 valid synonym-based questions, as opposed to the usual 10, due to the nature of the words selected by the students. The words they chose, with the one exception mentioned above, were in the text, but the students were not required to recall all the words during the questions, as is necessary when I dictate all the synonym sets. So, although the resultant data is not an apples-to-oranges comparison, perhaps it is akin to comparing a novella to a chapter.

The final data analyzed was of 15 synonym sets that were in the pre- and post-test. I provide details of the synonym sets selected in the section “Drawbacks of AI-Assisted Learner-Selected Synonym Sets” below.

My pre- post-tests reach an acceptable internal consistency with a Cronbach’s alpha of 0.80 or better each time. Therefore, the tests are reliable. To prevent interference caused by additional unknown words or the students being able to deduce the meaning from context, my pre- and post-tests use the following format.

Which word is a synonym of WORD?: 1 correct answer & 3 distracters

In assessing students’ receptive vocabulary knowledge, it is crucial to account for the possibility that correct answers on tests may not always reflect true understanding, as students

could select correct responses by chance. To address this issue, Nation and Webb (2010) emphasize the importance of using a receptive knowledge scale, such as the one developed by Waring, which allows researchers to more accurately gauge not only whether students know a word but also their level of certainty about that knowledge.

- 0: I do not know this word.
1. I have seen this word before, but I do not know its meaning.
2. I have seen this word before and know its meaning a little.
3. I know this word.

Nation & Webb, 2010, p. 299

This framework mitigates the issue of guessing correct answers and enables tracking of student confidence, offering insights into the depth of their vocabulary knowledge and their progression over time. Thus, I compare both student scores and their level of confidence on the pre- and post-tests.

Student Subjects

All 23 sophomores from a small liberal arts university in Japan who were eligible to participate in this research provided informed consent through an online test question. Their TOEIC scores varied widely, ranging from a low of 235 to a high of 730, with a median score of 407, indicating significant differences in English proficiency levels. Unfortunately, because of absences or failure to complete testing materials, only 13 students had sufficient data to be included in RQ 1 and 2, whereas all 23 are represented in RQ 3.

Drawbacks of AI-Assisted Learner-Selected Synonym Sets

While the use of AI-assisted learner-selected synonym sets has the potential to enhance student autonomy and engagement in vocabulary learning, this study has identified several drawbacks that warrant consideration and further research.

One significant drawback is the loss of study time incurred by allowing students to create their own synonym sets. In contrast to providing a pre-selected list of synonyms, which enables students to study immediately, creating personalized synonym sets requires a substantial time investment. In this study, students were given one week to find and select their own words, which effectively reduced study time compared to the teacher-provided condition. This time loss may have implications for the efficiency of vocabulary acquisition and the overall pace of the language learning curriculum.

Another drawback observed in this study was the difficulty in incorporating student-selected words into questions. While the AI-assisted approach aimed to empower students to choose words that applied to their individual learning needs, to maintain consistency among students, it was necessary to decide on 10 synonym sets to be used by the entire class. This negated the ability to focus on individual student needs. Unfortunately, I found that most of the words were not suitable for reading comprehension questions. I had to decrease the number of words included in the analysis because of issues like students selecting a word not mentioned

in the chapter and my poor predictions of student word choice on the pretest (see Methods). The original plan was to compare 10 synonym sets from the chapter using teacher-provided synonyms and 10 from the following chapter using AI-assisted student selected synonym. However, despite my best efforts to presuppose which words students may select from the text for the pre-test, the student-selected words were much more random. Therefore, this study only compares 5 words each from chapters 2 and 3 and the control words. This issue highlights the need for a balance between student autonomy and the practical constraints of assessment design.

Moreover, the AI-assisted learner-selected approach introduced a higher level of variability in the difficulty and complexity of the words chosen by students. While some students selected words that were appropriate for their current language proficiency, others chose words that were too simple or too advanced, which may have impacted their learning outcomes. Future iterations of this approach may benefit from providing students with clearer guidelines or parameters for word selection to ensure that their chosen synonyms can be effectively assessed within the given context.

Results

At the end of the semester, I conducted an anonymous survey to better understand how students felt about finding their own synonym sets. I could collect data from all 23 students. The following subsection previews the analysis of the pre- and post-test data by examining all 23 students' results on the tests given at the end of chapters 2 and 3. The last section is the results of the pre- and post-test for the 13 students who supplied sufficient data throughout the study.

Survey Results

AI-assisted synonym set creation was engaging, motivating, and beneficial for most students' vocabulary learning. The activity promoted learner autonomy, collaboration, and engagement in creating the list of words. While some students found aspects of the process challenging, the overall response was quite positive, with evidence of positive learning outcomes.

The students who found it challenging reported the following concerns: lack of confidence in finding appropriate synonyms, encountering difficult vocabulary, and the time-consuming nature of looking up words. Indeed, explaining how to use the AI and allowing students to use class time to work in groups took most of a 90-minute class period. However, only 8 of the 21 students preferred using synonym sets selected by a teacher. The other 13 students (62%) felt that creating their own synonym sets helped them learn and remember vocabulary better. They cited several reasons for their preferences, such as easier memorization, more specialized learning, and the benefits of looking up words themselves.

Several of the survey questions were on a 6-point Likert scale, with 6 representing the most positive response and 1 representing the most negative. Regarding engagement, most students found creating their own synonym sets to be quite interesting (average rating of 4.3

out of 6) and exciting compared to using pre-selected word sets (average 4.4). Additionally, most students reported being engaged in the process (average 4.6). Students also felt a strong sense of ownership and control over their vocabulary learning when creating their own sets (average 4.4). The most positive results were in group work and motivation: they found collaborating with classmates contributed considerably to their understanding of synonyms (average 4.8). However, this did not reflect on the use of AI. Whereas their motivation results do. They thought that creating their own synonym sets positively affected their motivation to learn new vocabulary (average 4.8). And many students answered they were likely to use this strategy in other courses or independent study (average 4.3), suggesting potential for transfer of learning.

Based on student motivation and their indication of usage, I followed up with a questionnaire after the summer vacation to see if students did indeed use AI to study vocabulary between semesters. The results were as follows: this paragraph is a reservation of space for the results of the survey to be conducted on September 30, the first class back after the vacation. Overall AI usage for language learning was fairly low among the students over the summer break, with most reporting no usage at all or only 1-5 times, suggesting that AI has not yet become a mainstream or habitual study tool for most of these learners. However, lack of knowledge or familiarity with how to use AI for studying was a top reason for not using it, pointing to the need for training and support to help students take advantage of AI tools effectively and address concerns about odd translations and inconsistent responses. Despite the low usage, many students felt AI had some positive influence on their motivation to learn and expressed openness to using AI more if they understood it better. Very few saw any negatives to using AI, though there were some reservations that it's no substitute for real human interaction.

These findings support the value of integrating AI tools like ChatGPT into vocabulary instruction to facilitate student-centered, exploratory learning. Several students voicing frustration in finding words in the text. Thus, I realize I need to provide clearer directions and guidelines. One student suggested having the teacher pre-select target vocabulary and allowing more time for the activity. One solution may be to give the students a suggested list of words from which to choose to see how this affects student perceptions. Future studies will reflect this.

For now, the survey results suggest students perceive generating their own synonym sets as an effective and valuable approach to vocabulary acquisition. The activity promotes engagement, motivation, collaboration, and a sense of ownership over their learning. While there are some challenges to address, especially in getting the students to adopt AI synonym searches into their independent study, the overall positive perceptions and potential of AI tools like ChatGPT into vocabulary instruction.

Chapter Test Results

At first glance, the average scores suggest students performed better with the synonym sets that were teacher-provided. For the chapter 2 test, all 23 students had an average score of 84.1%. And there was a sizeable drop to 65.7% on the chapter 3 test. This trend continues when comparing the percentage of students answering each question correctly. In chapter 2, over

80% of the students scored above 80%, whereas only 10 out of the 23 students (43%) achieved a score of 80% or higher on the chapter 3 test. There is much room for improvement in the effectiveness of the AI-assisted synonym set creation activity.

Another issue comes to light with the low proficiency students. On the chapter 2 test, only four students scored less than 50%. Those same students scored less than 50% on the chapter 3 test and only one other student joined their ranks. This would indicate that factors other than those being researched influenced their outcomes.

While the average scores suggest students performed better with the teacher-provided word list, the variable study duration and external disruptions make it challenging to draw definitive conclusions. The AI-assisted approach showed promise for some students, but grade inconsistencies and a decrease in student performance show it may not be universally effective.

Pre- and Post-tests

As mentioned in the methods section, I conducted a pre-test before both chapters to check student vocabulary acquisition. The timing for the post-test for chapter 3 was disrupted, thus, I decided to conduct a delayed post-test to assess long-term retention. The post-test was held on the first session of the reading strategies class after the summer vacation. The improvement in test scores for Teacher-Provided Synonym Sets (TPSS) was significantly higher than the control words, indicating that the activity effectively helped students learn and retain the target vocabulary. Although the improvement in test scores for the AI-Assisted Learner-Selected Synonym Sets (AALSS) was not statistically significant compared to the control words, it was still higher, suggesting that the activity may have had some positive impact on vocabulary acquisition for those words as well. Student confidence increased only slightly for the TPSS and dipped into the negative for the AALSS. And even more negative for the control words.

TABLE 1: Pre-test Post-test variance

score increase	control	0
	teacher-selected synonyms	32%
	student-selected synonyms	25%
confidence increase	control	-5%
	teacher-selected synonyms	6%
	student-selected synonyms	-2%

Additionally, I looked at the t-statistics between students' scores and confidence.

TABLE 2: t-statistics and P value

t-statistics	control	0	P	1
	teacher-selected synonyms	-3.384		0.028
	student-selected synonyms	-3.993		0.015

Conclusion

Learning synonyms can help students expand their vocabulary by using familiar words in sets that allow them to learn multiple words that convey similar meanings. Familiarity with synonyms can also aid in reading and listening comprehension. If a student encounters an unfamiliar word but knows its synonym, they can still grasp the meaning of the text or conversation. This is valuable in the context of language proficiency tests like the TOEIC, where understanding words in synonym sets can foster success.

Survey results indicated that students found AI-assisted learner-selected synonym sets (AALSS) engaging, motivating, and beneficial for vocabulary learning. The activity fostered learner autonomy, collaboration, and active word processing, with students appreciating the opportunity to take control of their learning. However, challenges were also noted, including difficulties in finding appropriate synonyms, time constraints, and vocabulary complexity. Despite these issues, 62% believed creating their own sets enhanced learning and retention and many felt that creating their own synonym sets boosted their motivation to learn new vocabulary.

The chapter test results, however, showed that students performed better with teacher-provided synonym sets. The chapter 2 test, using TPSS, had an average score of 84.1%, whereas there was a drop to 65.7% with AALSS being implemented for the chapter 3 test. The data also revealed that low-proficiency students consistently underperformed across both tests, showing external factors may have influenced their outcomes.

The post-test revealed the teacher-selected synonyms as having the better outcome. The larger improvement suggests that the teacher's expertise in selecting appropriate synonyms may have contributed to better student learning. However, it's important to note that both methods led to significant, albeit minor, improvements and the difference in p-values is not substantial. Therefore, AALSS shows promise for some students, but its effectiveness may not apply universally, requiring further refinement to address challenges and enhance learning outcomes. This variability may, in fact, underscore the importance of developing AI systems that can provide more personalized guidance and recommendations based on individual student needs and abilities.

Additionally, students expressed a strong likelihood of applying the AI strategy in other courses or independent study. In the results of the survey after the summer vacation, students expressed openness to using AI more if they understood it better. However, they felt AI is no substitute for real human interaction. Thus, I would recommend framing AI as a complementary tool, not a replacement for classroom instruction, to set appropriate expectations. Increasing awareness, training, and integration of AI into coursework could boost engagement and expand its utility for various language skills.

Despite the challenges encountered during the AALSS unit, this action research project yielded valuable insights. Students were impressed with the ease of obtaining context-valid synonyms using AI, demonstrating the potential of AI-assisted learning in vocabulary acquisition. While the study faced logistical challenges and limitations, it provides a foundation for further research into the effectiveness of AI-assisted vocabulary learning strategies. Future

studies should aim to refine the methodology, ensure closer monitoring of word selection, and allocate sufficient time for both the TPSS and AALSS units to enable a more balanced comparison.

This action research project was chaotic however, I have gained insights for more structured investigations into the potential of AI in language learning. By granting students greater autonomy and control over their vocabulary learning, this approach has the potential to foster increased motivation, engagement, and long-term retention.

Further Research

Vocabulary researchers should be precise in their terminology, especially avoiding blanket statements regarding vocabulary knowledge, and distinguish whether their research focuses on receptive or productive vocabulary knowledge, or vocabulary size or depth. I hope that studies on VLS will be clearer on student proficiency to allow researchers to better determine optimal stages for teaching aspects of vocabulary, such as synonyms or collocations. By doing so, we can advance our collective understanding of second language vocabulary acquisition and provide valuable insights for educators seeking to support their students' lexical development.

For my next stage, I am planning to explore ways to address the challenges with AI-assisted learning and investigate whether a hybrid approach, combining teacher guidance with learner autonomy. I hope more researchers embrace AI to help us pinpoint learning stages. Thus, we may allow students to become autonomous learners who can chart their own growth and use teachers as coaches to reach the next level of learning. This is a new paradigm that will require a shift away from teacher created tests with more AI-assisted attention given to individual students—getting there will take some creative solutions.

References

- Alharbi, A. M. (2015). Building vocabulary for language learning: Approach for ESL learners to study new vocabulary. *Journal of International Students*, 5 (4), 501-511. <https://www.ojed.org/index.php/jis/article/view/412>
- Alharbi, A. M. (2019). *Investigating sequential vocabulary learning strategies as a means of improving L2 vocabulary acquisition* [Doctoral dissertation, University of Memphis]. University of Memphis Digital Commons. <https://digitalcommons.memphis.edu/etd/2408>
- Arendt, H. (2018). *The human condition* (2nd ed.). University of Chicago Press. <https://doi.org/10.7208/chicago/9780226586748.001.0001>
- Aşık, A., Vural, A. Ş., & Akpınar, K. D. (2016). Lexical awareness and development through data driven learning: Attitudes and beliefs of EFL learners. *Journal of Education and Training Studies*, 4(3), 87-96. <http://dx.doi.org/10.11114/jets.v4i3.1223>
- Boers, F. (2021). *Evaluating Second Language Vocabulary and Grammar Instruction: A Synthesis of the Research on Teaching Words, Phrases, and Patterns*. New York, USA, Routledge.
- Britannica Dictionary. (n.d.). Find definitions & meanings of words. Retrieved September 22, 2024, from <https://www.britannica.com/dictionary>
- Charkova, D., & Charkova, K. (2018). Exploring the connection between second language vocabulary learning strategies and vocabulary knowledge. *International Online Journal of Education and Teaching (IOJET)*, 5(2), 235-250. <http://iojet.org/index.php/IOJET/article/view/317/232>

- Crompton, H., & Burke, D. (2023). Artificial intelligence in higher education: The state of the field. *International Journal of Educational Technology in Higher Education*, 20, Article 22. <https://doi.org/10.1186/s41239-023-00392-8>
- Educational Testing Service. (2024). *TOEIC® Listening and Reading: A complete practice test developed by ETS®*. https://siteresources.blob.core.windows.net/siteresourcescontainer/hauportal/media/root/orfeas/ets/toEIC/toEIC_sample-test_a4_64pg_v30.pdf
- Nevitt, C. A. (2023). Daily practice regimes for vocabulary and synonym acquisition and retention: Effects of including L1 meaning. *広島文教大学紀要 [Hiroshima Bunkyo University Bulletin]*, 58, 1-10. <https://doi.org/10.51095/kiyo.58.01>
- Read, J. (2004). Plumbing the depths: How should the construct of vocabulary knowledge be defined? In P. Bogaards & B. Laufer (Eds.), *Vocabulary in a second language: Selection, acquisition, and testing*. John Benjamins. <https://doi.org/10.1075/llt.10.15rea>
- Schmitt, N. (2010). Key issues in teaching and learning vocabulary. In Chacón-Beltrán, R., Abello-Contesse, C., and Torreblanca-López, M. (eds.), *Insights into Non-native Vocabulary Teaching and Learning*. Multilingual Matters.
- Webb, S. (2007). The Effects of Synonymy on Second-language Vocabulary Learning. *Reading in a Foreign Language*, 19 (2), 120-136.

—2024年9月24日 受理—