

AI in Language Education: Transforming the ESL Learning Experience

ABSTRACT

In Pakistan, EFL/ESL learners seek to improve their language skills through formal and informal means. Embracing AI, a key aspect of Industry 4.0, can empower the youth to become independent and efficient in language learning. This study aims to gauge the effectiveness of generative AI tools in learning vocabulary and evaluate ESL learners' user experiences and perceptions. Involving 12 students from a public sector university in Rawalpindi, this study has employed the case study approach with a triangulation design. The quantitative data is collected using pre and post-tests and the qualitative data is collected using semi-structured interviews. Between the pre-test and the post-test, the participants engaged with ChatGPT focusing on the 3 Cs of 21st-century skills i.e. communication, creativity, and critical thinking, following the Cognitive Taxonomy model by Anderson and Krathwohl (2001), for four weeks. The quantitative results (pre-test: $M = 18.08$, $SD = 6.598$ and post-test: $M = 31.33$, $SD = 6.946$) reveal a substantial difference of 13.25 between the mean scores which indicates a considerable growth in participants' vocabulary after the intervention. Whereas, qualitative results indicate positive experiences and perceptions of using ChatGPT for vocabulary learning, while challenges include response reliability, time management, and user suspicion of AI chatbot accuracy. Thus, the study concludes that ChatGPT and other similar AI-powered tools can be the panacea for language learning and teaching in the 21st-century world, given their appropriate and systematic use. This research can aid in developing effective teaching methodologies and integrating AI tools in language learning curricula, benefiting learners, researchers, and educators.

Keywords: ESL, Vocabulary, Generative AI, 21st-century skills, EdTech, ChatGPT, LLM

* National University of Science and Technology (NUST) Islamabad
aaminasyed15@gmail.com

** Fatima Jinnah Women University (FJWU) Rawalpindi
akifa.imtiaz@fjwu.edu.pk (corresponding author)

Introduction

The Fourth Industrial Revolution, also known as Industry 4.0 had a significant impact on education, requiring a new approach to education and training. As Farhat (2023) argued that as a key outcome of this industrial revolution, Education Technology (henceforth, EdTech) leverages digital tools to enhance teaching and learning processes, making education more adaptive, interactive, and aligned with the needs of the modern world. Artificial Intelligence (henceforth, AI) plays a crucial role in EdTech. According to De la Vall & Araya (2023), the application of AI in education (AIEd) had a positive impact on pedagogy through technical advancements and theoretical innovations. It has attracted considerable attention due to its numerous advantages for learners, such as personalized learning pathways, enhanced accessibility, increased motivation, support for active and self-directed learning, and access to a wide range of educational resources. For instance, it is being used to create adaptive testing systems that can adjust the difficulty of questions based on the student's responses which helps to identify knowledge gaps and provide targeted feedback to students. As with all other EFL/ESL learners, Shahzada et al. (2012) argue that, even after significant investments of time and resources, the majority of Pakistani students struggle with English language proficiency and their skills in second language/s are not well-developed. So, second language users constantly try to improve their language skills, especially vocabulary through formal and informal means.

With time, utilizing AI technologies for language learning is becoming common among autonomous language learners. As Woo and Choi (2021) pointed out learning and teaching of SLs have been revolutionized by a greater emphasis on individualized learning and rapid developments in AI as it provides immediate feedback on language skills, helping students improve pronunciation, grammar, and vocabulary. Moreover, as they include interactive roleplay, narrative, and simulated text messaging features, chatbots are useful tools for assisting students with the challenges of using a second language. One of them is ChatGPT and Paul et al. (2023) argue that AI-based tools like ChatGPT have revolutionised how customers think, act, and live in the post-pandemic world. Language translation, code generation, chatting, and content production are just a few of the many uses of ChatGPT. Interestingly, it performs all the tasks carried out by the previous AI tools individually. Thus, it has a great deal of potential to be utilised as a single, all-purpose tool to enhance English vocabulary instead of different AI tools for developing each type of vocabulary.

Hence, the present case study focuses on ascertaining the advantages and disadvantages of AI in ESL learning and teaching, as well as exploring

recommendations for Pakistani ESL learners to capitalize on the advantages and overcome the associated challenges.

Research Questions

The research questions for the study are:

- How effective is ChatGPT for learning vocabulary for undergraduate ESL Learners?
- What are the experiences and perceptions of the selected ESL learners about ChatGPT for learning vocabulary?

Significance of the Study

The present study explores the case of ChatGPT use for Online Informal Learning of English (OILE). Paul et al. (2023) noticed that ever since its launch, ChatGPT has generated a great deal of discussion regarding its primary characteristics, benefits, and potential downsides, Therefore, a study on this technology can contribute to the identification of specific methods for efficiently increasing English vocabulary and the development of solutions that are practical for learners. Since AI-powered bots are quite a new phenomenon there is a dearth of research on their role in language education in Pakistan. The present study is significant because it is a thorough multidimensional study of AI-based language learning practices and their impact on English vocabulary at the undergraduate level. Moreover, it contributes to the existing literature on language learning technologies and provides practical recommendations for learners, educators, and developers regarding the integration of AI-powered tools like ChatGPT in EFL/ESL classrooms and self-learning journeys. The findings also inform instructional strategies and support decision-making processes related to the use of such technologies for vocabulary learning. Although the present study gauged only ChatGPT for ESL learning, the findings can guide and facilitate those researchers who will study upcoming generative AI tools in the future, by employing similar study designs and methods.

Literature Review

AI tools are emerging technological mediums now being used in education in addition to several other fields, with language learning being an area on the verge of immensely benefitting from the rapid growth of the AI industry. Kuleto et al. (2021) discussed the impact of AI and Machine Learning (ML) on education, which they argue, have increased learning efficiency by aligning platforms and applications with learners' needs. They explored the potential and application areas of AI and ML in higher education through secondary research and a survey conducted among Serbian students and found that these technologies enhance learning by improving students' skills and facilitating collaborative learning.

Haristiani (2019), explored the potential of using chatbots as a language learning medium and found that chatbots can be effective as both language tutors and independent learning tools. The accessibility and flexibility of the chatbots increased learners' confidence as well as interest in using chatbots. Similarly, Woo and Choi (2021) conducted a systematic review of AI-based language learning tools which examined the impact of AI on the development of adaptive language learning tools and highlighted concerns regarding limited information and teacher preparation in utilizing these tools effectively. However, the users of the tools demonstrated improvements in language abilities and knowledge.

The incorporation of AI in vocabulary learning is also a potential area to be studied. Hsu et al. (2023) evaluated how young learners' vocabulary acquisition, self-regulation, and learning anxiety were affected by AI-supported image recognition technologies and self-regulated learning (SRL) and found that the group using AI-supported image recognition performed better in vocabulary acquisition and self-regulation development and exhibited lower learning anxiety compared to the control group. This way, it claimed that integrating AI technologies and SRL in language classrooms to enhance vocabulary learning outcomes can be fruitful.

Scope of ChatGPT as a Language-Learning Tool

ChatGPT 4, has garnered substantial media and tech industry interest. Cotton et al. (2023) explored the use of ChatGPT in academia and its impact on higher education. Enhanced student engagement, collaboration, and accessibility were identified as some of the benefits of using ChatGPT. Tilili et al. (2023) also explored the use of ChatGPT in the field of education, in three stages. The first stage examined the public discourse on social media, which generally showed positive attitudes towards using ChatGPT in educational settings along with some cautionary voices. The second stage focused on educational transformation, response quality, usefulness, and ethical considerations related to ChatGPT. The third stage investigated user experiences through ten educational scenarios, uncovering issues such as cheating, honesty, privacy concerns, misleading information, and manipulation. The study emphasized the need for further research to ensure the safe and responsible adoption of chatbots in education.

To be more precise, AI tools like ChatGPT have also garnered increasing attention in ESL learning worldwide. For instance, Kim, Shim, and Shim (2023) examined the use of ChatGPT in constructing and delivering business English writing courses for Korean ESL learners, finding that the tool could support task-based instruction and generate relevant course materials. Similarly, Fountoulakis (2023) investigated various AI-driven tools, including ChatGPT, in Greece and found notable improvements in language proficiency and intercultural communication. These international studies

highlight ChatGPT's adaptability to different linguistic and pedagogical contexts. However, while such research provides valuable insights into AI's global applications, there remains a significant gap in understanding how these tools perform in underrepresented contexts like Pakistan, where access, digital literacy, and localized pedagogy present distinct challenges. This study aims to address that gap by examining the use of ChatGPT for vocabulary development among Pakistani ESL learners, contributing both a regional perspective and practical insights to the global conversation on AI in language education.

Evolving EdTech and AI

The ever-evolving AI is likely to be a part and parcel of the EdTech landscape contributing to effective teaching and learning. For instance, *Grammarly* makes suggestions for improving grammar and punctuation after analysing written content using AI algorithms. An AI-powered language learning app, *Duolingo*, offers interactive activities and tailored feedback for people looking to improve their English. *Lingoda* uses speech recognition technology that is AI-driven to evaluate a user's pronunciation and speaking abilities. Another AI-powered study tool, *Quizlet*, uses flashcards, games, and quizzes to help students learn English vocabulary. An AI-based writing assistant, *Readable* improves material readability by examining sentence structure and word choice. However, Kannan and Munday (2018) noticed that while Massive Open Online Courses (MOOCs) initially (in the 2000s) overshadowed the influence of AI in language education, ongoing projects like Udacity, Khan Academy, EdX, and Coursera have incorporated AI. Moreover, a more developed form of AI i.e. conversational agents are simulated human-like interfaces utilised in Intelligent Tutoring Systems (ITS) for language learning. Bailey (2019) claimed that conversational agents can enhance learner participation and improve learning outcomes.

Research Methodology

The present study is an exploratory research with a case study approach. The research paradigm for the present study is based on a mixed-methods approach, combining both quantitative and qualitative methods. This approach was selected because it allowed for a comprehensive understanding of the effectiveness of ChatGPT for learning vocabulary, as well as the experiences and perceptions of the selected learners.

Research Design

Triangulation of design is applied to the present study. For this study, a triangulation design by Creswell and Plano-Clark (2007) is used. The following figure illustrates the design.

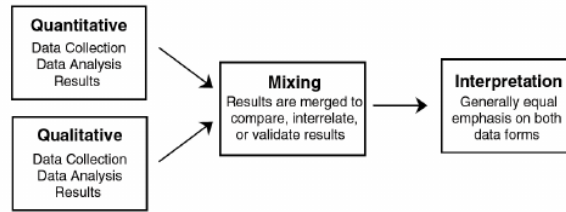


Figure 1. Triangulation design adopted from Creswell and Plano-Clark (2007)

The present research adopted the case study approach to explore and understand the effects of AI on the vocabulary of Pakistani ESL learners. By selecting a single case, which in the context of the present study was a group of ESL learners using the same AI tool, the research delved deeply into the intricacies of AI intervention's impact on English vocabulary learning. This allowed for a detailed investigation of the participants' experiences, perceptions, and challenges related to the use of ChatGPT, an AI-enabled generative tool.

Sample and Sampling Technique

For this research, twelve undergraduate ESL learners who were enrolled in different disciplines of Social Sciences at a public sector university were selected. A purposive sampling technique was used to select a diverse group of students based on their language proficiency and almost none or very little prior experience with AI-enabled language learning technologies. Initially, 21 students were selected but some of the participants did not continue their activities and left the research after informing us. All of them had a common reason behind leaving in between, and that was time management along with studies. So, only 12 successfully completed all the given tasks in the designated period and reached the post-test stage. In addition, semi-structured interviews were conducted with 4 participants.

Data Collection Tools

Following the data triangulation method, pre-test, post-test, and semi-structured interviews were used for data collection. For the test, common and trending 21st-century topics were selected so that the participants' vocabulary bank is improved in the areas which are crucial for global intelligence and world knowledge rather than too scientific or specialized areas. As, Chung and Nation (2003) suggest that, technical vocabulary - a set of words with specialised meanings that are specific to a particular subject area - is crucial in specialised texts, accounting for anywhere between 20 and 30 per cent of the text's running terms. Therefore, for the pre and post-tests, those vocabulary items were chosen which are the most significant and relevant words in the present times. Pre and post-tests were the same, they

had six sections each with a total of 50 points. All the sections employed different techniques to assess the vocabulary of the participants. Details of the test are described in Table 1 and Figure 2 below:

Table 1. Details of pre-test and post-test

Section	Category	No. of questions	Score
1	Definitions	15	15
2	Substitution	5	5
3	Synonym/Antonym	5	5
4	Contextual Explanation	2	10
5	Cloze Test	10	10
6	Idioms	5	5

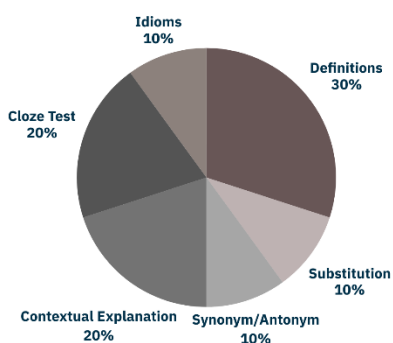


Figure 2. Pie chart on the Composition of pre-test and post-test in percentages

Moreover, the semi-structured interview had 12 main questions followed by a few follow-up questions tailored to participants' responses. These questions were designed to explore learners' perceptions, experiences, and challenges while using ChatGPT for vocabulary learning, focusing particularly on their engagement with AI-generated tasks, contextual vocabulary use, and the development of critical thinking. The questions addressed various dimensions of the learners' engagement, such as their initial expectations (Q2, Q3), perceived benefits and challenges (Q4, Q5, Q7), comparative

experiences with other digital tools (Q6), and their comfort level and interaction quality with the AI model (Q9). Several questions also focused on identifying specific features of ChatGPT that supported vocabulary learning (Q8), its impact on learner motivation (Q10), and the potential for integrating AI tools like ChatGPT into formal curricula (Q12). This way, by capturing both affective and cognitive aspects of learner experience, the interviews enabled a deeper understanding of how learners interacted with AI-based instruction. This qualitative component was added to provide rich, in-depth insights that could complement the quantitative pre- and post-test results, thereby justifying the use of a mixed-method design to better understand the impact of AI-based tools on language learning experiences.

Designing of AI-Based Vocabulary-Learning Activities

In the designing of the activities, a three-level approach was implemented to ensure contextual use of vocabulary among the participants to focus on communication, creativity, and critical thinking as these are the core 21st-century skills in education and learning. Regarding this, Wagner (2008) argues that in the 21st-century Conceptual Age, the immediate availability of information necessitates having critical thinking skills and being able to access and creatively evaluate information from many different sources. By integrating these three Cs into vocabulary learning, the rationale was to enable learners to progress from Lower-Order Thinking Skills (LOTS) to Higher-Order Thinking Skills (HOTS) (c.f Figure 3), following the revised Bloom's Taxonomy also known as Cognitive Taxonomy, updated by Anderson and Krathwohl (2001). As Singh and Marappan (2020) analyzed the use of HOTS in EFL/ESL learning and argued that use of HOTS aims to help the students to be independent, problem- solver and decision-makers in the future and they also claimed that the implementation of HOTS for ESL/EFL learning enables students to think critically through active learning. Having said that, the first level was designed so that they engaged in basic recall and understanding of vocabulary words (Remembering and Understanding). In the second level, they moved towards Applying and Analyzing stages, where the activities were related to reading texts incorporating new vocabulary, recognizing contextual meanings, and evaluating how words were used in different situations. In the third level, there was Evaluating and Creating stages, students were to write their responses using the learned vocabulary and critically reviewed their writing with ChatGPT's feedback. This cognitive shift made the students learn new words while being able to use them meaningfully and creatively in real-world contexts, reinforcing their vocabulary retention and linguistic competence. To reiterate, through these three-tiered activities; critical thinking was fostered, communication skills were enhanced, and creativity was developed by analyzing word meanings in different contexts, engaging in discussions, generating own writing pieces,

responding to prompts in diverse ways, and experimenting with vocabulary in unique contexts.

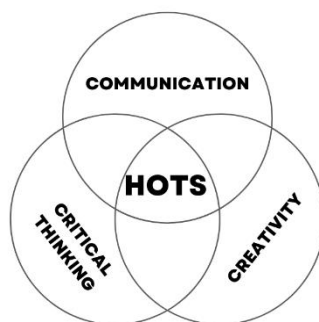


Figure 3. Relationship of the 3Cs and HOTS

Notably, ChatGPT is primarily not a language learning tool; so, considering the NLP capabilities, conversational interface, and personalized language practice feature of ChatGPT, the activity plan (c.f. Table 2) was carefully designed to transform it into an effective language learning tool. The first level of activities made use of ChatGPT's ability to understand and generate human-like responses in natural language. The second level was designed to focus on ChatGPT's ability to generate coherent and contextually relevant responses. Likewise, the third level with writing tasks, utilized ChatGPT's language generation capabilities to provide real-time support and guidance and focused on gauging ChatGPT as a writing assistant so that, as they write, participants could share their opinions and ideas, and ChatGPT offered feedback. This three-level approach aimed to provide a comprehensive learning experience that not only expanded participants' vocabulary knowledge but also fostered their creativity and critical thinking ability to use the vocabulary words in appropriate contexts (communication skills).

Table 2. Details of the activities performed using ChatGPT during the intervention

	Week 1	Week 2	Week 3	Week 4
Topics	1. Sports 2. Gender Studies	3. Environment 4. Water & Energy Crises	5. World Politics/IR 6. Conflict Resolution	7. Entrepreneurship 8. Idioms
	Activities			
1 hour	<ol style="list-style-type: none"> 1. Ask ChatGPT to generate a list of 15-20 advanced vocabulary and jargon related to the topic 2. Ask it to create a table of synonyms and antonyms of the words 3. Ask it to give an example of each word, how it will be used in real-life situations. 4. Ask for explanation of the examples. 5. Ask it to create 1 MCQ exercise to check your understanding. 			
0:45 hour	<ol style="list-style-type: none"> 1. Ask it to write a short essay on the topic using all those words in context and explain why a particular word was used. 2. Read the essay and understand how the words are used in the sentences and whether you fully understand the essay. 3. Give contextual explanation of the terms and ask it to review your response. 4. Ask for contextual explanations of specific terms in a tabular form. 			
1:15 hour	<ol style="list-style-type: none"> 1. Read the article provided to you and identify 10-15 specific terms related to the topic used in the article. Ask ChatGPT to explain any term/s which confuses you. 2. Now, write a 150-300 word extension or reaction to the article incorporating your total learned vocabulary on the topic. 3. Ask it to review your writing. Read the review thoroughly. 			

Students were engaged for a total of 6-8 hours per week. The activities were designed for a minimum of 6 hours, but considering variations in absorption and processing speeds of different learners, the time could exceed one to two hours.

Data Collection Procedure

After the participants signed informed consent, a pen-and-paper pre-test was conducted in person to record the learners' skills in English vocabulary before any intervention. The test was time-bound and completing the test took about 20-25 minutes. As the participants were not very familiar with using generative AI tools so a training session was held virtually on Zoom in which a detailed step-by-step demonstration was given to the participants on using ChatGPT and performing the activities for the given topics. After that, to examine the effects of the AI-based tool on the vocabulary of Pakistani ESL learners, there was an intervention of AI-based tasks using ChatGPT for four weeks. After the completion of the four weeks' intervention of learning vocabulary through ChatGPT-based activities, a pen and paper post-test was conducted in person to record the improvement in vocabulary, if any. After the results of pre and post-tests was examined using a Paired Sample T-Test, 4 participants were selected for the semi-structured interviews. Among them, 2 were the ones who showed the most improvement and 2 participants with

the least improvement in the post-test. These interviews were helpful in qualitatively analyzing various factors which may have affected the performance alongside the quantitative approach to their language skills.

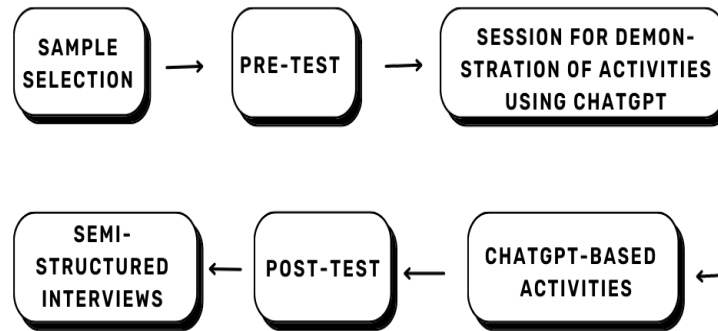


Figure 4. Flowchart of the data collection procedure

Ethical Considerations

The present research adhered to ethical guidelines and obtained informed consent from the participants. Confidentiality and anonymity of participants were ensured as the participants were coded as St1, St2... St12 for the tests and interviews and the data was stored securely. Before the commencement of the research, the students were provided with detailed information regarding the aim and objectives of the study, and they were given a chance to withdraw at any stage during the study.

Data Analysis and Results

Firstly, for the quantitative analysis, the scores from pre-tests and post-tests were analyzed through Paired Sample T-test. Secondly, for the qualitative analysis, the responses obtained from the interviews were thematically analyzed to validate the results of the quantitative analysis. Finally, findings were formulated by mixing the qualitative and quantitative results and making interpretations accordingly.

Analysis of Pre-test and Post-test

Table 3 and Figure 5 below show the pre-test and post-test scores and the differences between the two scores.

Table 3 Pre-test and post-test Scores

	Pre-test score out of 50	Post-test score out of 50	Difference in score
St 1	13	29	16
St 2	19	25	6
St 3	11	25	14
St 4	25	37	12
St 5	24	43	19
St 6	18	23	5
St 7	16	30	14
St 8	15	31	16
St 9	25	38	13
St 10	30	42	12
St 11	10	26	16
St 12	11	27	16

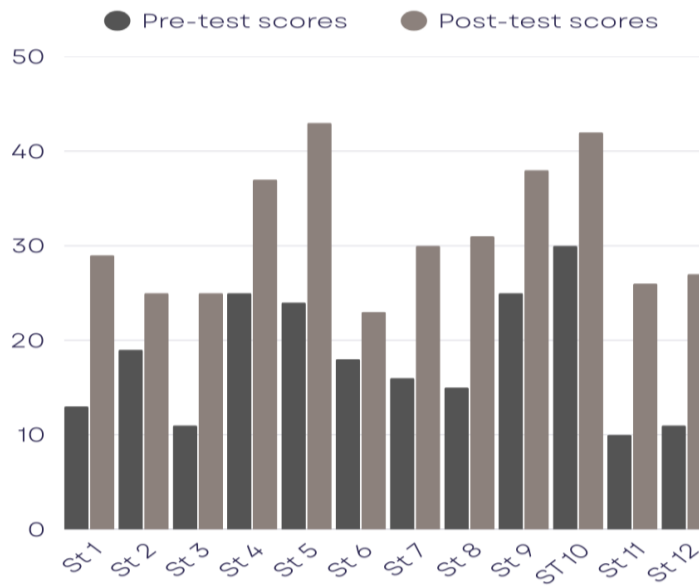


Figure 5. Bar chart displaying the differences in the pre-test and post-test scores

The collected results of participants' performance in the pre-test and the post-test were statistically analyzed through SPSS statistical software. To calculate the differences in the results of the participants, a Paired Sample T-test was run on the pre-test and post-test scores. The statistical values obtained from the Paired Sample t-test are presented below:

Table 4 Paired Sample Test

	Paired differences							
Pair 1	95% confidence interval of the difference					t	df	Sig. (2-tailed)
	Mean	Std. deviation	Std. error mean	Lower	Upper			
pretest results - posttest results	-13.250	4.137	1.194	-15.878	-10.622	-11.095	11	.000

From Table 4, the output of the SPSS Paired Sample T-test can be observed. Upon the analysis of the values of the pre-test and post-test scores through this test to evaluate the effectiveness of the tool quantitatively, the two-tailed paired sample T-test revealed that the participants' (N = 12) vocabulary bank increased from pre-intervention (M = 18.08, SD = 6.598) to post-intervention (M = 31.33, SD = 6.946); $t = -11.095$, $p < .000$, $df = 11$. Here, the t value measured how far the mean of the differences between two sets of data was from zero, divided by the standard deviation value (how much variation there is in the differences). Table 4 shows that the t value was a large number i.e., -11.095 and if the t value is large (or a large negative number) as in this case, then consequently the p -value will be small. As it was observed that in both cases, the p -value was less than 0.05 ($< .05$) therefore the difference between the scores before and after the intervention was highly significant. All in all, the results of the two-tailed Paired Sample T-test indicated that after the four-week intervention of AI through ChatGPT (c.f. Figure 6), there was a significant increase in the research participants' performance in the post-test, implying that the research participants improved their vocabulary bank on trending topics of the 21st-century world. This implication was corroborated by the participants' perceptions and evaluation of the selected AI tool's (i.e. ChatGPT in this case) effectiveness in their interview responses, where they admitted to having greatly benefitted from the tool and the overall intervention period.

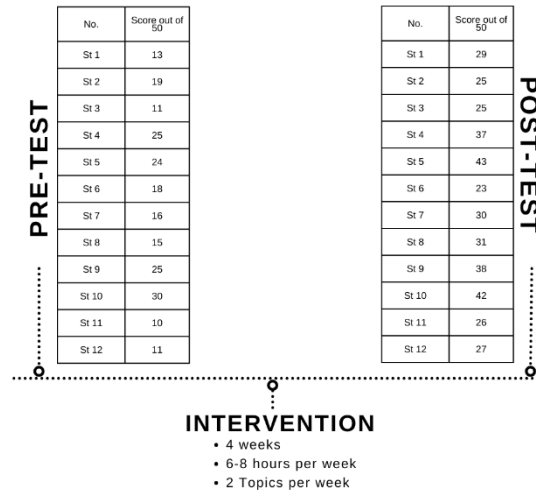


Figure 6. Timeline of Pre-test, Intervention period, and Post-test

Intervention of AI

For an effective intervention period, the participants were first provided with a detailed plan for a set of vocabulary tasks for each week, which they needed to complete using ChatGPT. The first level of the designed tasks encompassed five steps which provided a curated list of technical words of the particular area, expanding learners' vocabulary range; along with synonyms and antonyms in a tabular form promoting a deeper understanding of word relationships. At the same level, to enable the participants to grasp the practical application of the terms, there was a provision of real-life examples demonstrating their contextual use. Towards the end of the first level, any exercise of participants' choice was generated by ChatGPT to assess learners' understanding, promoting active recall and consolidation of the newly learned vocabulary. In the second level, there were more advanced tasks to deepen vocabulary learning such as asking ChatGPT to write a short essay on the topic using all the vocabulary words of the first level in context. This allowed the participants to observe, using their critical thinking, how the words are used in meaningful sentences and understand them in a larger context by focusing on the connotations, cultural nuances, and subtleties associated with specific vocabulary choices. Lastly, seeking contextual explanations for the terms which were challenging for the participants in a tabular form could aid in clarifying any misconceptions and reinforcing comprehension and retention of the vocabulary.

Moving on to the third and final level, it involved tasks such as reading a relevant article and identifying 10-15 technical vocabulary items of the area, to enhance vocabulary recognition and comprehension; as towards the

end of each week, we provided an article each, which was taken from the web, on the topics of that week so that the participants can self-evaluate whether what they had learned that week on the topic made them better able to comprehend a piece of writing on that topic. Even for their understanding of the articles, the participants used ChatGPT to recheck their understanding. Then, the subsequent task of writing a 150-300-word extension or reaction to the article, incorporating the learned vocabulary in the previous steps, strengthened vocabulary usage, critical thinking, and creative expression as it encouraged the participants to think creatively as they searched for appropriate words and phrases to convey their thoughts. As the final step of the level, the participants requested ChatGPT to review their writing and give feedback to help them refine their language skills and vocabulary application by thoroughly reading the review. This not only made them assess their progress after all the steps but also helped them identify areas for improvement. So, the participants did not only get inputs from ChatGPT but also gave their output on their learning and got that reviewed by the tool so that the effectiveness of ChatGPT as an instructor or learning partner could be gauged.

Besides, a few challenges arose during the intervention. One challenge was convincing the participants to perform the tasks properly and honestly. As the participants were from the junior-most semester, some of them were under pressure from projects and quizzes for the other courses they were taking; so, time management was a challenge for them which affected their availability to complete the tasks regularly. Therefore, to communicate the importance of their consistent participation to gather reliable data, every week, we had to emphasize on performing the activities regularly. However, reminders and regular check-ins helped keep most of the participants (N = 12) motivated and on track. Overall, it was observed that effective communication and consistent support from our side were crucial in addressing these problems and ensuring the successful implementation of the AI-based intervention for examining the effects on the vocabulary of Pakistani ESL learners.

Qualitative Feedback

According to the responses of the participants, three themes were deduced from the interview responses and on the basis of those themes generated after coding the responses, a qualitative analysis was conducted. The analysis following the three themes is as follows:

Theme 1: Motivations, Expectations, and Perceptions

The participants were asked about their motivations and goals behind taking part in the research which is based on using ChatGPT for learning vocabulary and they were also asked whether using ChatGPT as a vocabulary learning tool influenced their interest or motivation in improving their vocabulary.

Furthermore, they were asked to compare their expectations and perceptions with this experience when they decided to participate and when they actually used it. The theme of reasons to use AI tools was formulated based on their responses. One of the participants said: *"I always wanted to use good vocabulary and advanced words in my English writing such as essays, stories and paragraphs and this was my only motivation to be a part of the research based on using ChatGPT. Similarly, another participant revealed that: "Sometimes ago I had a habit of learning a word a day but as the time passed I somehow ditched this habit and then when you approached me and asked me that if I wanted to volunteer and participate in the research based on ChatGPT for learning vocabulary I thought to myself that it may be a good chance to rekindle my habit and so I instantly agreed to the offer and did not hesitate."*

In the responses, the participants mentioned *"curiosity"*, *"interest in improving vocabulary"* and *"possessing a good vocabulary bank"*. Curiosity and the hype associated with the latest tools were among the most common motivating factors behind the growing interest in chatbots. As Brandtzaeg & Følstad (2017) identified curiosity among other key motivational factors behind using chatbots as users view them as a novel phenomenon. Secondly, they also found productivity to be a major motivational factor that drives them to use LLM chatbots. Furthermore, the participants also mentioned the change in their interest in vocabulary learning after the intervention period. As one participant said: *"My vocabulary has improved by participating in this research and I found it very motivating for me to learn new vocabulary in the future because it really impacted my vocabulary."*

Theme 2: The Experience during the Intervention Period

The participants were also asked about their experience with the tool, its features, and its usability and to explain how comfortable they were in interacting with ChatGPT as a language model for vocabulary learning. They were also asked to elucidate their previous experiences of learning vocabulary using technological tools, if any, and compare them with the use of ChatGPT. In the responses, one of the participants said: *"It was a very overwhelming experience for me, overall it was very good. I tried different new things, and I learned about different new meanings and the use of ChatGPT in vocabulary. So, it was very beneficial for me so I think I would try it again in the coming semester."*

Likewise, in their responses, the participants also discussed their previous experiences with learning vocabulary independently using technological tools in comparison to their experience with ChatGPT and rated them. Interestingly, both, the participants who had no prior experience with technological tools for vocabulary learning, and the ones who had some prior experience, highly rated it. So, the positive experiences and perceptions

evident in the qualitative data analysis were in accordance with the results of the quantitative data analysis.

Theme 3: Opportunities and Challenges

As most of the responses from the participants were positive and they perceived the tool as beneficial, the participants were asked about the opportunities they saw that came with using AI-generated tools and they also discussed some challenges they faced. The responses showed that the participants benefitted from the tool's user-friendliness and good memory. However, the responses also identified a few challenges and limitations of the tool. For example, the time management along with their routine studies, the suspicion regarding the accuracy of ChatGPT's responses and its lack of knowledge of very recent real-life events hindered users from learning contextual vocabulary with real-life scenarios. The tool can also fall prey to misinformation available on the internet using which it formulates its responses, as Gordon Crovitz, reported in The New York Times article "Disinformation Researchers Raise Alarms About A.I. Chatbots" that ChatGPT is a potentially powerful tool for spreading misinformation that has ever been on the internet (Hsu & Thompson, Feb 2023).

Towards the end, considering the effectiveness and challenges related to the selected AI tool, ChatGPT, the participants were asked if they saw any opportunities related to ChatGPT as a language learning tool and if they would recommend it to others. They were also asked should AI-generated tools for language learning be formally incorporated in curricula and, if yes, then how and for what areas. All of them recommended ChatGPT to others *"...not just for vocabulary learning but for clearing any concepts or any doubts they have during their studies"*. Also, they supported the idea that *"...it can be formally incorporated in our curriculum but I would not suggest it for junior levels It should only be incorporated at the university level. Because I feel that if schools or colleges incorporate AI then the students will not brainstorm on things and might misuse it."* Another participant had a similar view, *"I would say that if it [ChatGPT/AI tools] is incorporated at the university level then students should be given awareness regarding it beforehand."*

Findings and Discussion

This section presents the findings of this research which have been interpreted from the discussion on the amalgamated results obtained from the qualitative and quantitative analyses. It was found through the quantitative analysis that the AI-based intervention using ChatGPT had a positive impact on the vocabulary development of the research participants. Likewise, the qualitative analysis revealed that the intervention period was a fruitful experience for the

participants. Imtiaz, Umer & Akhtar (2021) also found that despite facing a few challenges, students exhibited a favorable perception of utilizing online materials and that positive perception had a positive impact on their learning practices too, as the students found online language learning materials effective in augmenting their learning experiences.

The analysis of the two sets of collected data unveiled the powers of ChatGPT. The findings declare ChatGPT as an effective language learning tool that can also be used for language teaching as using the latest AI technologies can provide learners with personalized and interactive language practice, enabling them to expand their vocabulary and enhance their overall language skills. One important point to note is that ChatGPT, by default, is not a language-learning tool. However, for the present research, we leveraged ChatGPT to teach vocabulary, and the post-test results of the participants demonstrate how effectively this tool can be exploited for language learning.

The participants' accounts of their experience of using the tool revealed different factors behind the effectiveness of ChatGPT. The unique conversational style, good memory, and interactive nature of tabular and bulleted short answers to the queries make the tool user-friendly and conducive to learning. However, it was also found that the tool poses some challenges in terms of the reliability of the information and learners who used the tool in informal settings independently had constraints like time management and trusting the information at times. These challenges and constraints can hinder their learning but it can be overcome with external human guidance or cross-checking from other authentic online sources in case of unavailability of human resources.

The findings further demonstrate that the factors of innovation associated with recent AI tools and the user-friendly and interactive nature of ChatGPT act as motivating forces for not only using the tool but also boosting independent learning in users to enhance their educational performance as well as general knowledge. Moreover, despite the challenges and hindrances they endured, the participants voted in favour of the incorporation of AI tools into the curriculum. Nevertheless, incorporating the latest AI technologies in formal and institutional settings would work best if incorporation is in terms of utilizing the tool during classroom activities and lecture proceedings to enhance the effect of teaching.

Notably, these findings resonate with emerging international research on the integration of AI tools in second language learning. For example, Kim, Shim, and Shim (2023) explored the capabilities of ChatGPT in designing and delivering a business English writing course for a Korean ESL learner persona using the Task-Based Language Teaching (TBLT) framework. They examined how ChatGPT could generate course content and interact as a teaching assistant, concluding that while the tool was not without limitations,

it showed significant promise in supporting structured language learning. Likewise, Fountoulakis (2023) conducted a large-scale mixed-methods study involving 200 participants in Greece, evaluating the impact of AI tools like Duolingo, Rosetta Stone, and ChatGPT on both language proficiency and intercultural communication. The study found that AI tools contributed to notable improvements in learners' practical language skills and cultural awareness, especially when used for personalized and interactive learning experiences. These studies strengthen the validity of the present research's findings by confirming that AI tools, including ChatGPT, can meaningfully enhance ESL learning in diverse contexts. However, while their focus was either on structured course design or broader intercultural outcomes, the present study contributes a unique perspective by emphasizing vocabulary development and critical thinking in a developing country setting. It also brings attention to the local challenges: digital literacy, learner motivation, and time constraints, that can shape the effectiveness of AI-based interventions. Thus, this research not only supports but also extends the global discourse by offering insights from a less-represented context, affirming the pedagogical potential of ChatGPT for ESL learners beyond technologically advanced or resource-rich environments.

Suggestions for Future Research

The realm of AI is evolving at such a pace that new tools and technologies are emerging at a remarkable pace with the launch of one tool after the other having varying functions and capabilities. This makes it essential for researchers to stay abreast of these developments and conduct investigations to gauge the role of AI in educational technology. So, as we look ahead, future research will be driven by the question of whether to combine or replace previous technological approaches with AI technologies in face-to-face contexts to maximize learning outcomes or not and why. Therefore;

- Future researchers can conduct a study to examine the long-term effects of using or incorporating AI-powered generative tools like ChatGPT to assess ESL/EFL learners' vocabulary retention and continued growth beyond a specific period of intervention. This would lead to comparing short-term vocabulary knowledge and long-term retention to determine if the observed improvements in the post-intervention performance are sustainable over time.
- A replicative study would be fruitful with a larger and more diverse sample (with participants from diverse academic disciplines) and a generative AI tool other than ChatGPT. Research samples can be with varying language proficiency levels, and educational stages/age groups to assess the applicability of AI-based intervention across different learner populations; so that the generalizability of the findings can be ensured.

- As per the scope and word limit of the present study, this research has focused only on a single micro skill of the language, i.e. vocabulary. So, future researchers can investigate the effectiveness of AI-powered generative tools in learning the other three micro-skills of English, or the macro skills for that matter, to add to deducing the effectiveness of incorporating AI technologies in ESL/EFL learning.

Conclusion

Industry 4.0 has made it necessary for everyone living in the present world to have adequate global knowledge and English communication skills to be able to have meaningful and effective conversations in local as well as global contexts; nevertheless, in Pakistan, English is an FL and an SL, so the learners' English skills are usually not so advanced and they try to improve them through formal and informal means. As, with advancing and evolving EdTech and AI, Pakistan needs to embrace these developments of Industry 4.0 and renew its language learning pedagogies through the incorporation of the latest tech tools while coping with the challenges it posits. Thus, the study concludes on the note that ChatGPT and other similar AI-powered generative tools, can be the panacea for learning and teaching languages in the fast-paced 21st-century world; as this tool can not only act as a flexible teacher but also as a learning partner, making learning enjoyable, independent and economical, provided it is used appropriately and methodically. This will benefit not only the educators and traditional learners but also motivate independent learners with its interactive abilities to take charge of their learning.

References

- Bailey, D. (2019). Chatbots as conversational agents in the context of language learning. *Proceedings of the Fourth Industrial Revolution and Education*, 32–41.
- Brandtzaeg, P.B., Følstad, A. (2017). Why People Use Chatbots. In *Internet Science: 4th International Conference, INSCI 2017, Thessaloniki, Greece, November 22–24, 2017, Proceedings 4* (pp. 377–392). Springer International Publishing. https://doi.org/10.1007/978-3-319-70284-1_30
- Chung, M., Nation, P. (2003) Technical vocabulary in specialised texts. *Reading in a Foreign Language*, 15(2), 103–116.

- Cotton, D. R., Cotton, P. A., & Shipway, J. R. (2023). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. *Innovations in Education and Teaching International*, 1–12. <https://doi.org/10.1080/14703297.2023.2190148>
- Creswell, J. W. (2012). Educational research: Planning, conducting, and evaluating quantitative and qualitative research (4th ed.). Boston, MA: Pearson.
- Creswell, J. W., & Clark, V. L. P. (2007). *Designing and conducting mixed methods research*. Sage Publications, Inc.
- De la Vall, R. R. F., & Araya, F. G. (2023). Exploring the benefits and challenges of AI-language learning tools. *International Journal of Social Sciences and Humanities Invention*, 10(01), 7569–7576.
- Farhat, L. (2023). From E-learning to Ed-tech; A Case Study of India's Road to Education 4.0. Focus.
- Fountoulakis, M. S. (2024). Evaluating the impact of AI tools on language proficiency and intercultural communication in second language education. *International Journal of Second and Foreign Language Education*, 3(1), 12–26.
- Haristiani, N. (2019, November). Artificial Intelligence (AI) chatbot as language learning medium: An inquiry. In *Journal of Physics: Conference Series* (Vol. 1387, No. 1, p. 012020). IOP Publishing.
- Hsu, T. C., Chang, C., & Jen, T. H. (2023). Artificial Intelligence image recognition using self-regulation learning strategies: effects on vocabulary acquisition, learning anxiety, and learning behaviours of English language learners. *Interactive Learning Environments*, 1–19. <https://doi.org/10.1080/10494820.2023.2165508>
- Imtiaz, A., Umer, S., & Akhtar, Y. (2021). Technology in action: perceptions & practices of students in using online language learning material at higher education level. *Psychol. Educ. J*, 58, 11207-11218.
- Kim, S., Shim, J., & Shim, J. (2023). A study on the utilization of OpenAI ChatGPT as a second language learning tool. *Journal of Multimedia Information System*, 10(1), 79–88.

- Kuleto, V., Ilić, M., Dumangiu, M., Ranković, M., Martins, O. M., Păun, D., & Mihoreanu, L. (2021). Exploring opportunities and challenges of artificial intelligence and machine learning in higher education institutions. *Sustainability*, 13(18), 10424. <https://doi.org/10.3390/su131810424>
- Majhanovich, S. (2014, March 13). Neo-liberalism, globalization, language policy and practice issues in the Asia-Pacific Region. *Asia Pacific Journal of Education*, 168–183.
- Panda, M. and Mohanty (2015). Multilingual education in South Asia: The burden of the double divide. *Handbook of Bilingual and Multilingual Education*. New Jersey: Wiley-Blackwell, pp. 542–553.
- Paul, J., Ueno, A., & Dennis, C. (2023). ChatGPT and consumers: benefits, pitfalls and future research Agenda. *International Journal of Consumer Studies*. <https://doi.org/10.1111/ijcs.12928>
- Schmitt, N. (Ed.). (2013). *An introduction to applied linguistics*. Routledge.
- Shahzada, G., Ghazi, R., & Khan, A., (2012). Causes of Rising Failure of the Students' in the Subject of English at Secondary Level. *Mediterranean Journal of Social Sciences*. 3 (1), 605–611.
- Singh, C. K. S., & Marappan, P. (2020). A review of research on the importance of higher order thinking skills (HOTS) in teaching english language. *Journal of Critical Reviews*, 7(8), 740–747.
- Tlili, A., Shehata, B., Adarkwah, M. A., Bozkurt, A., Hickey, D. T., Huang, R., & Agyemang, B. (2023). What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education? *Smart Learning Environments*, 10(1), 15.
- Wagner, T. (2008). *The global achievement gap*. New York: Basic Books.
- Woo, J. H., & Choi, H. (2021). Systematic review for AI-based language learning tools. *arXiv preprint arXiv:2111.04455*.