

Empowering ESL Teachers with Generative AI

Peter Snashall

Faculty of Liberal Arts, Huachiew Chalermprakiet University

*Email : eslflow@gmail.com

Abstract

This research paper investigates the potential of generative AI technology, specifically large language models (LLMs), in enhancing English as a Second Language (ESL) instruction by offering teachers an innovative tool to create engaging and dynamic classroom activities. In addition to emphasizing the rapid development of LLMs and machine learning in language education, the study also underscores the necessity for teachers to acquire the skill of prompt engineering to harness the full benefits of AI. The methodology combines qualitative and quantitative survey data collected from ESL teachers who evaluated the usefulness and relevance of AI-generated prompts created by advanced LLMs for various ESL exercises, such as role-plays, vocabulary quizzes, and listening tasks. The research also examines the potential benefits and challenges associated with integrating generative AI, including LLMs and machine learning, into language teaching. The survey results reveal a positive response from teachers, indicating that generative AI and LLMs can enhance the quality and variety of ESL activities and save teachers valuable preparation time. By using AI to creatively generate exercises and mastering prompt engineering (writing instructions for AI), teachers can increase student engagement, promote language acquisition, and accommodate diverse topics, learning styles, and proficiency levels. However, concerns regarding the accuracy of AI-generated content are also highlighted, emphasizing the importance of proofreading, critical evaluation and customization of AI-generated materials by teachers. The paper concludes by discussing the implications of these findings for ESL instruction and future research in the field of language education, LLMs, and AI technology. Recommendations are provided for best practices in leveraging generative AI tools, including LLMs, and developing prompt engineering skills to maximize their potential in ESL teaching while mitigating potential risks and challenges. This research contributes to the growing body of literature on the intersection of AI, LLMs, and language education, and its findings hold promise for enhancing the teaching and learning experience in ESL classrooms.

Keywords : generative AI, AI prompts, teacher development, teacher materials, teacher efficiency, prompt engineering

1. Introduction

Teachers are overworked and underpaid, and they have to deal with a lot of administrative paperwork. They also have to prepare for classes with inadequate resources. They have to grade assignments and exams, learn online skills, and attend meetings. Not only that, they have to deal with overcrowded classrooms and/or huge classes online. My thesis is that the new forms of Generative AI emerging now can provide invaluable assistance to teachers. It can help them to create customized courses for specific subjects and topics. It can help teachers to create quizzes and tests that are more effective in assessing student understanding. It can help teachers create interesting teaching materials. All these tasks can be done much more quickly. Consequently, teachers can manage their own time more effectively. In short, AI can help teachers be more creative, more efficient and effective in nearly every task they perform. Teachers who take advantage of AI have the opportunity to become highly valued professionals.

In fact, it is already being suggested that the ability to talk to AI could be the most important skill of the century (Warzel, 2023).

Generative AI is the artificial ability to do certain tasks faster and more efficiently than is possible with human intelligence. The goal of generative AI in education is to provide students with learning experiences that are tailored to their individual needs and abilities. Generative AI can be used by teachers to create customized lesson plans, assess student progress, and provide feedback. The benefits of using Generative AI in education include increased student engagement and motivation, improved academic performance, and reduced workload for teachers. Some potential challenges of using generative AI in education include the potential for bias and discrimination, the need for data privacy and security, and the need for teacher training. Generative AI can be used to improve teaching in the future by providing more personalized and effective instruction, reducing the workload for teachers, and improving communication between teachers and students.

The rapid evolution of artificial intelligence (AI), particularly through generative models such as GPT-4, introduces novel opportunities for educators. However, the path to such empowerment is not without its hurdles. Although AI models have markedly improved in recent years, challenges persist. Lu et al. (2023) underscore the necessity of appreciating not just the capabilities but also the limitations of these models, emphasizing the importance of understanding what prompts they can effectively process. AI's prowess, though impressive, is still marred by various errors, from grammatical and usage inaccuracies, through incoherent responses and self-contradictions, to factual errors and violations of common sense understanding (Dou et al., 2022). Therefore, while leveraging the power of AI in the realm of ESL teaching shows immense promise, critical scrutiny of AI systems and continuous development to overcome their shortcomings are imperative to ensure their reliability and applicability.

2. Objectives

1. Investigate the potential of generative AI technology, specifically large language models (LLMs), in enhancing ESL instruction by providing teachers with a tool to create engaging and dynamic classroom activities.
2. Explore the necessity for ESL teachers to acquire the skill of prompt engineering in order to harness the full benefits of AI-generated content for language education.
3. Analyze qualitative and quantitative survey data from ESL teachers who evaluated the usefulness and relevance of AI-generated prompts created by advanced LLMs for various ESL exercises.
4. Examine the potential benefits and challenges associated with integrating generative AI, including LLMs and machine learning, into language teaching, such as improved engagement, language acquisition, and personalized learning opportunities, as well as accuracy and cultural appropriateness concerns.
5. Evaluate the implications of these findings for ESL instruction, the field of language education, and the future research on LLMs and AI technology, discussing how they may contribute to the enhancement of the teaching and learning experience in ESL classrooms.

6. Provide recommendations for best practices in leveraging generative AI tools, including LLMs, and developing prompt engineering skills to maximize their potential in ESL teaching while mitigating potential risks and challenges.

7. Finally, to answer the research question: How can generative AI technology, specifically large language models, be utilized to empower ESL teachers, and what are the necessary skills, potential benefits, challenges, and best practices associated with integrating this technology into language instruction?

3. Materials and methods

This study aims to investigate the effectiveness of generative AI in empowering ESL teachers by providing them with innovative and useful teaching materials. In order to examine the efficacy of AI-generated prompts in producing valuable teaching resources, we employed a survey-based approach utilizing Google Forms. This section provides a detailed overview of the methodology followed in this research.

1. Participants

The participants in this study were a group of ESL teachers with varying levels of teaching experience, educational backgrounds, and teaching contexts. A total of 5 ESL teachers were randomly selected from an online database of educators to ensure a representative sample. The participants were informed about the purpose of the study and were ensured that their responses would remain confidential and be used solely for research purposes.

2. Materials

Six AI-generated prompts were created using a generative AI model, specifically designed to produce educational resources for ESL teachers. These prompts were designed to cover a range of language skills, including grammar, vocabulary, reading comprehension, and speaking activities. Each prompt was designed/engineered to create a corresponding AI-generated exercise or activity.

3. Survey Design

An online survey was created using Google Forms to collect data from the participants. The survey consisted of three main sections:

a) AI-generated Prompt Evaluation: In this section, participants were asked to evaluate the effectiveness of each of the six AI-generated prompts in producing useful teaching materials. A 10-point Likert scale was used, ranging from 1 (Strongly Disagree) to 10 (Strongly Agree). Participants were also asked to provide qualitative feedback in the form of comments for each prompt, elaborating on their evaluations.

b) Overall Impressions: The final section of the survey asked participants to provide their general thoughts on the potential of generative AI to empower ESL teachers and to suggest any improvements or further research directions.

4. Data Collection

The online survey link was distributed to selected ESL teachers via email, and they were given one week to complete the survey. A follow-up email was sent to remind the participants a week before the deadline. After the data collection period, the responses were exported from Google Forms and analyzed.

5. Data Analysis

Quantitative data from the Likert scale evaluations were analyzed using descriptive statistics to determine the overall effectiveness of the AI-generated prompts. Qualitative feedback from the comments section was analyzed using thematic analysis to identify recurring themes and patterns in the participants' perceptions of the prompts. This combined approach allowed for a more comprehensive understanding of the potential of generative AI in empowering ESL teachers.

4. Results

The survey for the usefulness of AI for teachers consisted of six prompts for ChatGPT (OpenAI). The prompts are an example of prompt engineering. They have been created through acquired AI skills, trial and error testing, and a deep knowledge and experience of the subject area, using OpenAI's ChatGPT-4 (a paid subscription service). Consequently, people who used Chat GPT-3 for the survey would have had inferior results. ChatGPT-4 is a superior application and many of the problems encountered in ChatGPT-3 have been eliminated. Previous research has noted some of the errors that can be found in GPT-3 and sometimes in GPT-4. This can be most easily understood in the infographic below (Dou et al., 2022).

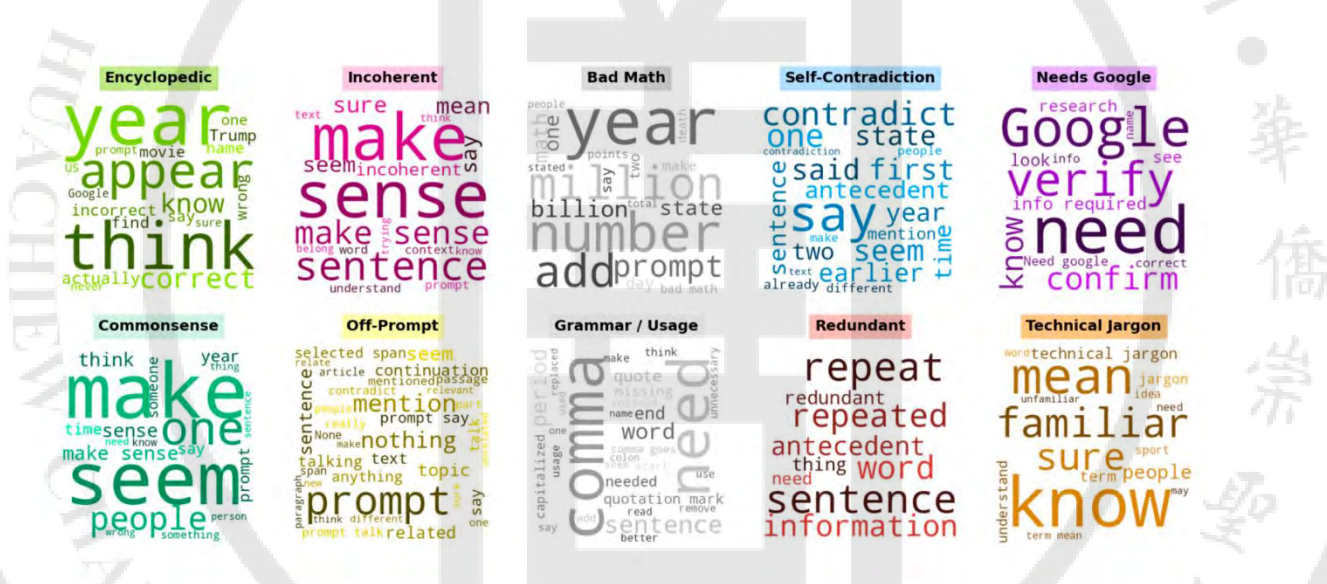


Figure 1: Common unigrams and bi-grams showing error type.

The prompts below were assessed for their usefulness from 1-10 as can be seen in following the charts and comments.

Table 1

	1	2	3	4	5	6	7	8	9	10
Prompt 1: Create a 10 line conversation for making a doctor's appointment between a receptionist and a patient as an English as a second language role play:								✓	✓	✓
Prompt 2: Make a create-your-own conversation exercise with a word bank for English as a second language students studying how to express opinions.								✓✓	✓	

	1	2	3	4	5	6	7	8	9	10
<p>Prompt 3: Create gap fill sentences with multiple choice answers for each of the items below, A/B/C:</p> <ol style="list-style-type: none"> 1. good service 2. pay employees well 3. reduce prices 4. good location 5. provide snacks 6. promotions 7. open 24 hours 						✓	✓		✓	
<p>Prompt 4: Create exercises for English as a second language students studying the topic “taste and nutrition” including: a ranking/prioritizing exercise; an opinion statement agree/disagree exercise; and a sentence starter complete-the-sentence with your own ideas exercise.</p>								✓	✓	
<p>Prompt 5: Create exercises for English as a second language students studying the environment including: a classifying exercise; a collocation exercise; and a fill-in-the-blank sentence exercise with a word bank.</p>								✓	✓	
<p>Prompt 6: Create an English as a second language conversations skills exercise, with 5 short gap fill conversations with a word bank, between students with different names for English as a second language students studying the topic: making decisions.</p>						✓	✓		✓	

Prompt 1: This prompt achieved a median rating of 9 in the survey, showing that all teachers thought it was impressive. The usefulness of the prompt in Table 1 lies in its focus on a common and essential task, making a doctor's appointment, which students might need to perform in their day-to-day lives. Furthermore, as the conversation is limited to 10 lines, it saves time and allows teachers to efficiently cover multiple topics within a single lesson. This brevity also encourages students to practice and memorize the dialogue, ultimately leading to better retention and application of the language. Overall, this prompt supports the development of functional English skills and promotes language acquisition through contextualized and meaningful practice.

Prompt 2: This prompt achieved a median rating of 8.5 in the survey. It creates an activity that not only encourages students to practice their vocabulary and grammatical skills, but it also encourages critical thinking and self-expression. By engaging in this exercise, students learn to construct sentences and articulate their thoughts effectively, while adapting to different conversational contexts. Additionally, the word bank serves as a valuable resource that helps students enrich their vocabulary and gain familiarity with various expressions used to convey opinions. As ESL teachers aim to develop their students' language proficiency, this exercise promotes active participation, collaboration, and confidence in speaking English, ultimately leading to a more comprehensive and meaningful learning experience.

Prompt 3: Prompt 3 focuses on the creation of multiple choice exercises. This prompt achieved a median rating of 7.6 in the survey. Some teachers weren't as impressed with this prompt. Perhaps, not all teachers like using multiple choice, or perhaps it requires more proofreading and editing. Or perhaps the teachers lack prompt engineering skills. This prompt can be extremely useful for English as a second language (ESL) teachers as it streamlines the process of generating engaging and effective learning materials. By providing an efficient framework for creating multiple exercises (questions or gap fill format), this prompt allows teachers to save valuable time and energy that can be redirected towards tailoring lessons to their students' needs, addressing individual challenges, and cultivating an interactive classroom environment. Additionally, the multiple-choice aspect of this format promotes a deeper understanding of the language, as students are required to not only identify the missing word but also choose the most appropriate option from a set of possibilities. This engages critical thinking and contextual understanding, further enhancing their language acquisition process. Overall, this prompt serves as a practical tool for ESL educators, simplifying the task of crafting educational resources and enabling a more effective learning experience.

Prompt 4: This prompt achieved an impressive median rating of 8.6 in the survey. Prompt 4 serves as an invaluable tool for English as a second language (ESL) teachers, as it provides a formula for the creation of diverse and engaging exercises centered around any topic. By incorporating a ranking/prioritizing exercise, an opinion statement agree/disagree exercise, and a sentence starter complete-the-sentence exercise, this prompt enables instructors to develop a comprehensive lesson plan that simultaneously targets various language skills. As students work through these activities, they have the opportunity to practice reading, writing, speaking, and critical thinking skills, all while expanding their vocabulary and understanding of the topic. Moreover, the prompt's multifaceted nature allows teachers to efficiently generate multiple exercises at once, saving valuable preparation time and ensuring that students receive a well-rounded learning experience.

Prompt 5: This prompt also achieved an impressive median rating of 8.3 in the survey. Prompt 5, which requests the creation of exercises for English as a second language (ESL) students studying a specific topic (the environment, in this case), can be particularly useful for ESL teachers as it adds variety to their lessons. By integrating different types of exercises, such as classifying, collocation, and fill-in-the-blank activities with a word bank, teachers can cater to diverse learning styles and keep students engaged. Furthermore, these exercises promote a well-rounded understanding of the subject matter by focusing on both vocabulary development and context comprehension. Overall, implementing this prompt in the classroom supports a dynamic, comprehensive, and relevant learning experience for ESL students.

Prompt 6: This prompt achieved a median rating of 7.6 in the survey. The rating is a little bit lower, perhaps because of the complexity of the prompt. Most people in the survey were using GPT3. I think more complex prompts require GPT4. This prompt (6) emphasizes the ability of AI to create extremely customized teaching materials. Focusing on conversation skills exercises with gap fill activities, the prompt enables teachers to craft engaging scenarios that reflect authentic interactions between students with different names. This diversity in names promotes inclusivity and helps students identify with the characters in the conversations. The prompt's emphasis on the topic of "making decisions" fosters critical thinking skills and relevant language acquisition, further enhancing the learning experience. By designing exercises that cater to

students' individual linguistic levels and interests, ESL teachers can significantly improve their students' communicative competence and overall language proficiency.

5. Discussion

Artificial intelligence is continuously improving, and the analysis of open-ended generative models such as GPT-4 sheds light on the limitations and strengths of these AI systems. Lu et al. (2023) emphasize the importance of understanding the capabilities of these models by analyzing the prompts they can and cannot handle. By categorizing prompts by their constraints, researchers can create clear bounds on model capabilities, which ultimately contributes to model explainability and points to future directions for improvement.

One example of AI improvement is evident in the transition from GPT-3 to GPT-4. GPT-4 exhibits enhanced performance and capabilities, particularly in tasks such as natural language understanding, text generation, question-answering, summarization, and translation (OpenAI). This improvement is due in part to a larger model size and more diverse training data, as well as better handling of constraints in prompts, such as document-type, structural, stylistic, and subject constraints (Lu et al., 2023).

Despite these improvements, AI models are still prone to errors. Dou et al. (2022) outline various error types found in AI-generated text, such as grammar and usage errors, off-prompt responses, redundancy, self-contradiction, and incoherence. Factual errors, such as bad math, encyclopedic inaccuracies, and violations of commonsense understanding, also persist. These errors reveal the ongoing need for scrutiny and development in AI systems to ensure they provide accurate and coherent information.

Nonetheless, artificial intelligence, particularly in the realm of open-ended generative models like GPT-4, is improving through a more fine-grained understanding of model capabilities and limitations (Lu et al., 2023). This understanding is crucial in driving explainability and providing direction for future advancements in AI technology. While progress has been made, it is important to continue addressing the errors and shortcomings of AI systems to ensure their reliability and usefulness in real-world applications (Dou et al., 2022).

One way of looking at the improvement in AI and large language models can be seen in the growing public awareness of its abilities. For example, Steve Johnson wrote in the *New York Times* in April 2022:

OpenAI's GPT-3 and other neural nets can now write original prose with mind-boggling fluency — a development that could have profound implications for the future.

And in an article in *The Atlantic Magazine*, writer Charlie Warzel discusses how prompt engineering, the skill of writing effective instructions for AI models, is becoming an essential 21st-century skill. Teachers are recognizing the potential of prompt engineering in their classrooms as well. He gives the example of Ethan Mollick, a professor at the University of Pennsylvania, who has adapted his syllabus since the release of ChatGPT to incorporate this skill. He asks his students to imagine ChatGPT as a student and teach it by prompting it to write an essay about a specific class concept. The students must then refine the AI's essay, much like a professor during office hours, until it demonstrates sufficient mastery of the subject. This exercise not only helps students learn by explaining, but it also teaches them the valuable skill of

writing effective prompts. As AI integration in various aspects of life continues to grow, prompt engineering will likely become a crucial skill for teachers and students alike.

Another article demonstrating the rapid advances of AI recently appeared in the technology magazine *Wired* in March 2022. Clive Thompson described the experience of a coder, Aboukhadijeh, when using the AI engine Copilot:

He discovered that it had other impressive tricks; it could even understand commands he wrote in basic English. If he simply typed into his code editor “Write a function that capitalizes every word in a document,” Copilot would assemble that code all by itself. He’d check to make sure it didn’t have errors; sometimes it did.

It is helpful to read technology journalists, coders and computer scientists, to keep up with how AI can help teachers more easily create engaging teaching materials (and assist them in many other tasks). Just as these journalists and coders have been experimenting with using generative AI capabilities for coding and language tasks, it can be used for creating teaching activities.

6. Conclusion

In conclusion, the potential for Generative AI to revolutionize ESL teaching and empower educators is significant, as evidenced by its capabilities in personalization, automation, interaction, and creativity. It provides a unique opportunity to create tailored educational content, enhance student engagement, and reduce administrative burdens on teachers.

However, it is paramount for educators to familiarize themselves with prompt engineering, to harness the potential of AI effectively. This skill will allow them to develop customized and specialized courses, automate various manual processes, quickly respond to student queries, and create diverse teaching materials in various formats. Consequently, educators will be able to cater to the specific needs of specific groups of students, as well as individual students, and ensure that the materials are tailored to their students' abilities. And as AI technology continues to expand its capabilities, teachers will be able to easily incorporate more diverse materials including visual aids, audio and video into their teaching activities. Of course, teachers must also remain vigilant and maintain an active role in the process, proofreading AI-generated content to mitigate inaccuracies or misinformation.

Future research on AI should focus on understanding the technology's ability to generate more diverse and accurate educational content, whilst addressing concerns over data privacy, security, and potential bias. Prompt engineering, is one aspect of AI that will undoubtedly become a focus of a lot of research. Future developments in the area should include training programs to equip teachers with the necessary AI skills, thus creating a future where the power of AI is harnessed efficiently and ethically for the betterment of ESL education. By bridging the gap between technology and pedagogy, we can democratize access to quality language education, enable more inclusive classrooms, and ultimately, create a more globally connected, linguistically diverse society. The fusion of AI and education holds promise, but it must be navigated wisely to ensure that the human element of teaching is preserved, and that educational outcomes are prioritized above all. As we tread this path, the collective goal should be leveraging AI not to replace educators, but to augment their capabilities and allow them to focus more on what they do best – inspiring and nurturing the minds of tomorrow.

7. References

- Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D. M., Wu, J., Winter, C., ... Amodei, D. (2020). Language models are few-shot learners. arXiv preprint arXiv:2005.14165.
- Cope, B., Kalantzis, M., & Sears, D. (2020). Artificial intelligence for education: Knowledge and its assessment in AI-enabled learning ecologies. Pages 1229-1245.
- Dou, Y., Forbes, M., Koncel-Kedziorski, R., Smith, N. A., & Choi, Y. (2022). Is GPT-3 text indistinguishable from human text? Scarecrow: A framework for scrutinizing machine text. arXiv preprint arXiv:2107.01294v3.
- Johnson, Steven A.I. Is Mastering Language. Should We Trust What It Says? (15th April, 2022) *New York Times Magazine*
- Lu, A., Zhang, H., Zhang, Y., Wang, X., & Yang, D. (2023). Bounding the capabilities of large language models in open text generation with prompt constraints. Georgia Institute of Technology, Shanghai Jiao Tong University, Google, Stanford University.
- Mullen, C. A. (2017). *Artificial intelligence in education: Past, present, and future*. In C. A. Mullen (Ed.), *The Routledge international handbook of artificial intelligence in education* (pp. 3-21). New York, NY: Routledge.
- OpenAI. (n.d.). OpenAI: Discovering and enacting the path to safe artificial general intelligence. Retrieved from <https://openai.com/>
- Pelachaud, C., & Poggi, I. (Eds.). (2016). *Computational modeling of emotions and social signals*. Cham, Switzerland: Springer.
- Snashall, P. (2023). AI prompt survey [Unpublished raw data, Google Forms]. Huachiew Chalermprakiat University.
- Thompson, Clive, A.I. is My Copilot, (April,2022) *Wired Magazine*
- Warzel, C. (2023, February 9). The most important job skill of this century: Your work future could depend on how well you can talk to AI. *The Atlantic Magazine*.
- Wei, J., Wang, X., Schuurmans, D., Bosma, M., Ichter, B., Xia, F., Chi, E., Le, Q. V., & Zhou, D. (2022). Chain-of-thought prompting elicits reasoning in large language models. In *Advances in Neural Information Processing Systems 35 (NeurIPS 2022)*.
- Woodie, A. (2023, February 14). Prompt Engineer: The next hot job in AI. *Datanami*. <https://www.datanami.com/2023/02/14/prompt-engineer-the-next-hot-job-in-ai/>

9. Appendices

Appendix A: Creating a conversation



Create a 10 line conversation between a pharmacist and a customer who is buying aspirin.



Pharmacist: Hi, how can I help you today?

Customer: I need to buy some aspirin.

Pharmacist: Do you have a prescription for that?

Customer: No, I don't. Is that a problem?

Pharmacist: No, it's not a problem. Aspirin is available over-the-counter. How many do you need?

Customer: Just a bottle of 100 tablets, please.

Pharmacist: Alright, let me grab that for you. Anything else you need?

Customer: No, that's all. Thanks.

Pharmacist: Here you go. The total comes to \$5.99.

Customer: Great, thank you.

Pharmacist: You're welcome. Have a good day.