

# LARGE LANGUAGE MODEL-BASED ARTIFICIAL INTELLIGENCE IN THE LANGUAGE CLASSROOM: PRACTICAL IDEAS FOR TEACHING

by **Euan Bonner**

euan.bonner @ gmail.com

**Ryan Lege**

ryan.frederick.lege @ gmail.com

and **Erin Frazier**

efrazier.research @ gmail.com

Kanda University of International Studies

Chiba, Japan

## **Abstract**

Large Language Models (LLMs) are a powerful type of Artificial Intelligence (AI) that simulates how humans organize language and are able to interpret, predict, and generate text. This allows for contextual understanding of natural human language which enables the LLM to understand conversational human input and respond in a natural manner. Recent examples of this, such as the Generative Pre-Trained Transformer (GPT) model, popularized by OpenAI's web application, ChatGPT, are able to complete an astounding variety of tasks when provided with simple language input. For education, LLMs can alleviate teacher curriculum and grading workloads and even perform specific tasks such as generating creative ideas for activities. Specifically for language learning, LLMs can draw on their immense corpus of language content to generate learner-centric materials to aid teachers in delivering targeted, personalized language instruction. The aim of this paper is to provide the reader with examples of how LLMs can be utilized for materials development, classroom activities, and providing feedback. After giving specific examples and explanations, the paper will conclude with a discussion of how this technology can provide teachers with new innovative ways to streamline the teaching process to focus on learner needs.

**Keywords:** ChatGPT; Artificial Intelligence; Large Language Models; language education; CALL

## **1. Introduction**

Artificial Intelligence (AI) is a concept that educators should already be aware of and many may be interested in knowing how it will affect their jobs in the future (Marche, 2022). Many schools and universities have already taken advantage of Educational Data Mining, a tool that uses AI to process massive volumes of data to better understand students and the context in which they learn. It has been used to analyze the class attendance and assignment submissions of entire institutions

and provide support to students at risk of dropping out of school (Tsai et al., 2020). Numerous papers have also been written about its eventual role in helping teachers with curriculum planning, grading, student management, and more (Pokrivackova, 2019; Settles et al., 2020). AI has even been utilized in creating intelligent tutoring systems (ITS) that support students by providing personalized feedback and guidance (Kochmar et al., 2022). Shukla et al. (2019) remark that their ITS system was “not only capable of providing immediate intelligibility assessment, but also tracking the learner’s experience, which in long term can aid in improving the retention of the learning” (p. 64).

Within the field of foreign language learning, AI-powered tools have assisted students in accomplishing more with increasingly capable natural language processing and contextual understanding in conversation (Lu, 2018). AI tools such as Intelligent Personal Assistants (IPAs), including Amazon Alexa, Google Assistant, or Apple Siri, can understand various accents and interpret meaning from non-standard language production (Moussalli & Cardoso, 2020). This has allowed for their use in creative activities where students create and interact with their own personalized voice recognition systems used as tour guides (Frazier et al., 2020). Similarly, AI conversation partners have also been used in creative language learning activities (Cai et al., 2021, Fryer et al., 2017). Machine Translation (MT), using AI services such as Google Translate and DeepL, has long been evaluated for its affordances for language learning (Deng & Yu, 2022; Lee, 2019). Briggs (2018) revealed that most students use MT tools in their education, even if they are aware of their accuracy limitations. Lee (2019) investigated the use of MT in writing and showed that MT assisted students in decreasing the number of lexical and grammatical errors present in their work while improving their overall revision skills. Commercial applications such as Grammarly have also used the affordances of AI to assist language learners in improving the output of their foreign language writing.

However, a new kind of AI, Large Language Models (LLMs), has emerged as perhaps the most powerful tool that is affecting language teaching and learning today. LLMs have been trained on billions of human-generated texts that allows them to predict the next word in a given sequence, allowing it to generate coherent and contextually appropriate text. The ability of LLMs to interpret human written requests and predict and generate a natural human-like response, like student essays, has led many media outlets and educational institutions to predict enormous changes to existing educational paradigms (Roose, 2023; Gillani, 2023).

This paper will introduce the power and current affordances of LLMs and explore how they can reduce the amount of time teachers spend on classroom preparation and student feedback. For this paper, we utilized OpenAI’s ChatGPT (<https://chat.openai.com/>), but the principles

shared here could be applied to any LLM service. ChatGPT is a service where once users have created an account they are presented with a text input field. Users can then write *prompts* for the AI to complete using its LLM, called Generative Pre-Trained Transformer (GPT). Prompts can be written in natural, conversational language posed either as a question or incomplete statement, which ChatGPT will attempt to answer or complete. ChatGPT also features a built-in memory allowing users and the AI to reference previous parts of the conversation.

Prior to the release of ChatGPT, OpenAI's GPT Playground (<https://platform.openai.com/playground/>) was available in closed beta (see Figure 1). Over the course of 2020 and 2021, we investigated applying this technology to education by generating ideas for its potential use and creating textual prompts. These prompts were submitted to the AI and then the generated responses were evaluated for reliability and validity. Some of these prompts underwent numerous iterations and adjustments before the AI could generate accurate and useful responses consistently. The release of ChatGPT in late 2022 greatly simplified the way that humans can interact with LLMs, removing many of the barriers that prevented this technology from being used by the general public (see Figure 2). Using ChatGPT, we were able to simplify our workflow previously used in OpenAI's GPT Playground and use ChatGPT to complete the following tasks: *Summarize and level texts for learners, Automatically correct grammar and sentence mechanics, Compose narrative writing prompts, Create presentation notes, Generate lesson ideas and Level texts for testing or reading practices*. However, before we can understand how the AI accomplished this and provide details about how teachers can take advantage of these capabilities, we must first understand what AI really is, what LLMs are, and how they function in the context of being a teaching and learning aid.

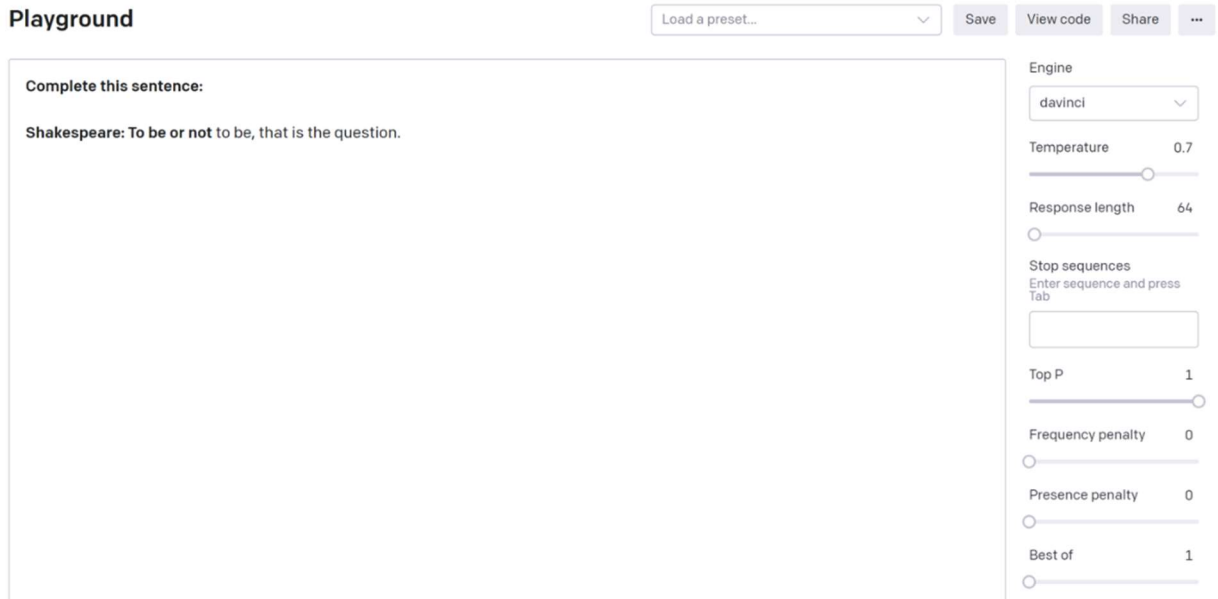


Figure 1. OpenAI's Playground command console

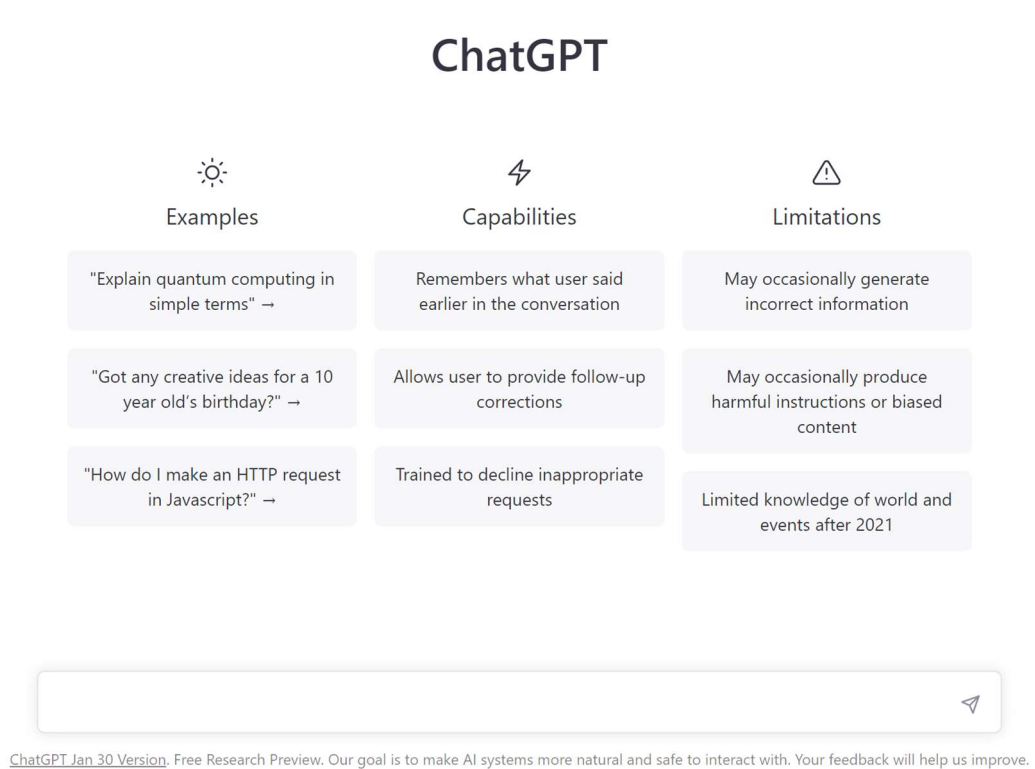


Figure 2. OpenAI's ChatGPT interface

## 2. Literature review

### 2.1. Thinking Machines

Alan Turing, the famous British computer scientist attributed as the father of computing and artificial intelligence, once stated in his seminal paper, “Computing Machinery and Intelligence,” that by the end of the 20th century, there would be no controversy in stating “that one will be able to speak of machines thinking without expecting to be contradicted” (Turing, 1950, p. 442). Nowadays, we offload an enormous amount of “thinking power” onto computers. Everything from calculating taxes through a smartphone calculator to using warehouses of supercomputers to calculate global weather patterns is a kind of computational thinking.

Turing, in addition to believing that computers could think, also felt that computers would ultimately start to think in the same way as humans, to the point where the difference between talking with a human or with a computer becomes indistinguishable. This type of thinking, one that mimics aspects of human intelligence and behavioral traits, is called Artificial Intelligence (AI). Distinct from the thought process used to calculate taxes or weather patterns, AI attempts to understand an aspect of human interaction and mimic it as best it can. In our everyday lives, AI is increasingly prevalent: Intelligent Personal Assistants (IPAs) listen to our speech and fulfill our requests and Automated Driving Systems (ADS) use cameras and sensors to process data about the driving environment to safely navigate to a destination.

### 2.2. Machine Learning

Modern AI, like those found in IPAs or in ADSs, is made possible by machine learning. Machine learning helps AIs improve their intelligence and performance by processing incredibly large amounts of data and being trained to identify patterns. As Holmes et al. (2019) state, “machine learning may be considered a three-step process (analyze data, build a model, undertake an action) that is continuously iterated (the outcomes of the action generate new data, which in turn amends the model, which in turn causes a new action). It is in this sense that the machine is learning” (p. 18). Over time, an AI becomes very adept at identifying patterns correctly, whether it is an IPA transcribing speech in a crowded environment, or an ADS identifying a stop sign while driving.

In education, machine learning AI services such as *Amazon Web Services* offer machine learning educational systems to institutions and claim to “spot struggling students earlier and take action to improve success and retention” (Machine Learning in Education, 2021). In foreign language education, machine learning has already been used to a great effect in automatic language translation (Niño, 2020). The Google Translate app can aid in automatic translation

merely by pointing a smartphone camera at a foreign language sign and translating it in a way that visually matches the fonts and colors of the original image (see Figure 3).



Figure 3. Google Translate

Recent research points to even greater possibilities in the near future. A publication by Google Research and the Indian Institute of Technology Kharagpur (Lahiri et al., 2021) focuses on automatically dubbing videos into additional languages. The AI not only provides an audio track in another language but crafts a voice that mimics the vocal character of the original speaker and modifies the video image so that mouth movements match the new language audio stream. This technology will likely be integrated into streaming platforms like YouTube, providing the same content in multiple languages without human input. This, of course, provides additional support for language learning by making authentic content more accessible.

### **2.3. Large Language Models**

Finally, we arrive at neural networks, a method of machine learning that has great promise and potential for teaching and learning. Generally speaking, machine learning AIs train with vast amounts of data, stored in categories, and tagged with content features very logically. Neural networks, on the other hand, organize this training data in a much more human-like fashion, categorizing information not only by its features but also by its relationship with other information. Neural networks, as suggested by the word “neural”, organize information by the same principles we believe the human brain is organized, as a vast mesh network of connections. The human brain is made up of billions of neurons, all interconnected, and it is through the combination and strengthening of these connections that memories are formed and learning takes

place (Holmes et al., 2019). Hence, a neural network is able to naturally form contextual relationships among immense amounts of information. Large Language Models (LLMs), a kind of neural network, are able to interpret natural human input by drawing from their own human-like memory, searching for the topic and related topics that best apply. If an LLM were to be asked to write an email to one's boss about getting time off from work, it would create a template drawn from its own understanding of an email's common features, then add its own creative sentences and key components derived from what it knows of human requests, polite interactions, and work culture.

The ability of LLMs to accomplish complex context-dependent tasks that previously, no matter how banal, needed a human touch is arguably their greatest asset. LLMs allow for the offloading of more mundane tasks (Pokrivcakova, 2019), allowing humans to focus on the social, creative aspects of work. Good instructional design starts by identifying learners' needs and knowledge deficiencies, organizing this information, and designing learning experiences for the students (Brown & Green, 2019). When education is tailor-made for the learner, it is most effective; but due to temporal, budgetary, or practical constraints, education is often relegated to assembly-line, cookie-cutter style approaches. If educators could overcome the hurdles mentioned above, educational quality would inevitably rise, as "by customizing the content of every student's curriculum, each student's learning could be optimized to help them in areas where they need it the most" (Arslan et al., 2021, p. 35). However, these hurdles are not that easily vaulted, as pursuing individualized education would necessitate exponentially increasing the teacher workforce to deal with this demanding task. Here is where LLMs can turn an impossibility into a reality. They can support the teacher by processing vast amounts of information about the students and their learning process, then use this information to support the creation of adaptive learning environments that are catered to the needs of individual learners.

### **3. Practical applications for teaching and learning**

Before going into each example use case of an LLM, it is important to understand how they receive and process input and generate output. Typically, computer programs receive instructions written in a language that is designed to be read by machines and not humans. This is not the case for LLMs. Services based on LLMs are designed to understand and process the same kind of instructions that a human would give to another human. Therefore, in order to have a service like ChatGPT generate a desired response, users need to provide the LLM with enough information for it to pick up on the context and continue to develop it. This also helps avoid



issues with LLMs, which are prone to faking information when their neural network database does not contain a clear answer (Summers, 2023).

The following section provides practical ideas for using ChatGPT to support teachers with their lesson preparation and student feedback. Each of the following examples will provide background information, provide a prompt to be submitted to ChatGPT, an example response, and an analysis of the output. These prompts can be modified to meet a teacher's particular learning context and need. Due to the nature of LLMs, everytime a prompt is submitted, there will be some variation in the response provided. Therefore, teachers are encouraged to submit the prompt multiple times and pick the response that is best for them.

### 3.1. Summarizing text in level appropriate language

The ability to summarize a text is often a skill practiced in a language learning context. LLMs are able to quickly and efficiently do this task to give learners not only a model of a well-written summary, but also provide learners with level-appropriate text.

Table 1. Example of summarizing text in level-appropriate language

---

| <b>Input</b>  |
|---|
| <p>My CEFR A2 student asked me what this passage means:</p> <p>Online options in higher education have expanded at an increasing rate over the last 10 years, from entirely online courses, to hybrid courses, to the greater prevalence of online components in traditional in-person courses. This growing trend will increase opportunities for non-traditional students to meet their educational needs (Nakamura, 2017), but at what cost to the quality of that education? How can institutions transition or modify materials to online settings, possibly toward a hybridized environment? In an effort to determine the efficacy of entirely online lessons vs. in-person lessons in the language-teaching field, this study gathered information from approximately 180 learners in a language-focused university in Japan about their perceptions of language learning in both online &amp; traditional settings. The study aimed to establish whether online, asynchronous versions of core curricula, integrated with in-person lessons, could offer comparable educational value. The findings suggest that although online lessons were viewed mostly positively overall, they were not deemed as valuable as the traditional in-class format. Increasing opportunities to interact in an online language learning environment seems to be an important consideration.</p> <p>I summarized it for him, in plain language a CEFR A2 can understand:</p> |

---



---

### Output (AI response)

---

People have started studying online more in the last 10 years, and there are different types of online classes. Some classes are all online, some are a mix of online and in-person, and some have online parts added to regular in-person classes. This is good for some students, but people are worried about the quality of education. A study was done to compare online language classes to in-person language classes. The study found that online classes were liked, but not as good as in-person classes. To make online classes better, students need more chances to talk with others online.

---

As the example above shows, the AI is able to break down an academic abstract into an easily digestible text for learners at a particular level, in this case, CEFR A2. When asking for a summary, one must be very specific about the level of text they wish the AI to produce. Specific educational level terminology must be used in order to extract accurate level-appropriate summaries (e.g., 2nd grader, 12th grader, CEFR A2, TOEFL 500). This capability of the AI could also be used to check student summarization skills by comparing their summaries to what the AI generates.

### 3.2. Correcting grammar and mechanics

One of the most menial, time-consuming tasks of language teaching is correction of grammar mistakes. Grammar corrections are rule-based, basically binary in nature, either right or wrong; therefore the perfect target for correction by AI. LLMs not only have an enormous corpus of language to draw on, but they are also aware of what is considered correct language. In the example below shown in Table 2, the “Standard American English” text is used to prompt the AI into changing student language samples riddled with grammar, spelling, and punctuation errors into correct text.

Table 2. Example of correcting grammar and mechanics

---

### Input

---

You are a professional English teacher who is good at grammar. You received the following sentences from your students. Correct the grammar and mechanics.

1. **Original:** Next, put sugar, salt and vinegar in a small bowl to make the vinegar for sushi.
  2. **Original:** First, cut the chicken and vegetables about the size of the meal easily.
  3. **Original:** First, put the oatmeal and water in a bowl.
  4. **Original:** First, Mix pork meat and ?ed vegetables . Pour some olive oil in a pan and heat
-

---

**Output (AI response)**

---

1. **Revised:** Next, put sugar, salt, and vinegar in a small bowl to make the vinegar for sushi.
  2. **Revised:** First, cut the chicken and vegetables to a size that is easily manageable for a meal.
  3. **Revised:** First, put the oatmeal and water into a bowl.
  4. **Revised:** First, mix the pork meat with chopped vegetables. Then, pour some olive oil into a pan and heat it.
- 

The AI is able to complete a holistic correction of the text, correcting not only grammar, but also incorrect use of punctuation or spacing. While a rather impressive result, it is important to remember that the AI's corrections are a result of it consulting its database of billions of examples of real-world English usage. As an LLM, the AI has formed connections between these examples, including their grammar and mechanical features and based on that, determined whether each specific grammatical and mechanical feature identified in the student's text conforms to them. So for grammar purists of the prescriptive perspective, the AI may not identify all of the errors in accordance with traditional grammar books. Rather, it is making corrections based on language as it is actually used, taking a more descriptive approach that may deviate away from rules taught by teachers. Furthermore, one needs to consider that the grammar model adopted by the AI may exclude world Englishes as proper language models.

### **3.3. Generating narrative writing prompts**

In writing courses, prompts are employed for a variety of purposes, giving students both a focus and catalyst for the writing process. Sometimes, as is the case for most standardized test writing prompts, these are carefully crafted to ensure that student writing adheres to a strict genre convention. In other cases, unexpected, offbeat prompts are used as an impetus for creative writing tasks. The AI can operate as a prompt generator, creating novel prompts that match the purpose of the writing task. Compared to random prompt generators, which essentially pull content from an existing list, LLMs can generate bespoke writing prompts that fit the instructional needs of a task. In the example presented below, the AI is given some core story elements such as information about the characters, setting, and plot, along with an example of how to convert this information into a story introduction. The result is a unique, creative prompt based on the information given (see Table 3).

Table 3. Example of generating narrative writing prompts

---

|  |
|--|
| <b>Input</b>   |
| Create a short story introduction based on the following information:<br><b>Character:</b> Lacy<br><b>Occupation:</b> English Teacher<br><b>Setting:</b> Waikiki beach<br><b>Action:</b> Learning to surf  |
| <b>Output (AI response)</b>  |
| Lacy, an English teacher, stood on the sandy shores of Waikiki beach, watching the waves crash against the shore. She had always been fascinated by the ocean, but never dared to venture into the water. Today, however, she was determined to conquer her fear and learn how to surf. With a determined look on her face, she approached her instructor, ready to take on the waves and embrace the thrill of surfing. |

---

As above, you can see that the AI created a logical story based on the input. Language learners could then use this prompt as the basis for creating a story in a narrative genre unit of their writing course. Rather than the teacher spending a considerable amount of time thinking up and writing down numerous individual creative writing prompts or simply resorting to sharing a single prompt with all students, the AI can quickly generate bespoke creative prompts for as many students as needed. Students are then free to practice their narrative genre writing skills, knowing that when they share their stories with others in follow-up activities, no two stories will be alike. The content and format of the input could be modified for many different genre conventions or to include student-generated content. In our testing, the AI consistently produced unique, usable prompts. These kinds of tasks, ones that require the generation of many ideas within a desired structure, are well suited for the creativity of LLMs.

### 3.4. Creating presentation notes

LLMs have the ability to parse text and identify main ideas quickly. This is relatively simple, especially in texts like presentation scripts that use signposting words and phrases. Presentations are often a key part of education, frequently with a focus put on both content and presentation skills. While there are many factors contributing to either the success or failure of the students, one key contributing factor is students' use of notes or prompts. Reading verbatim from a script is often the kindling for an uninspiring presentation. LLMs can simply convert student scripts

into presentation notes or outlines to guide learners so that they do not need to rely on verbatim scripts. In fact, any text can be broken down into its core components by the AI (see Table 4).

Table 4. Example of creating presentation notes

---

| <b>Input</b>   |
|--|
| <p><b>Script:</b> Today I'm going to talk about the history of data storage. Early computers initially used punch cards as a way to store data. Data was stored by punching in a hole in a specific column. The modern concept of a hard disk was first introduced in 1956 by IBM in the form of their computer called RAMAC. The system used massive, 24-inch diameter platters to store data. They only held 5 megabytes of data! Eventually, the 3.5 inch hard drive was invented, allowing for more data to be stored on small devices. Next, the 2.5 inch size drive was invented. Now, much of our data is stored on solid state drives that don't use spinning disks at all, storing data on computer chips.</p> <p>Summarize in bullet point form:</p> |
| <b>Output (AI response)</b>  |
| <ul style="list-style-type: none"><li>• Early computers used punch cards for data storage</li><li>• IBM introduced the concept of a hard disk in 1956 with their computer RAMAC</li><li>• RAMAC used 24-inch platters that held 5 MB of data</li><li>• 3.5-inch hard drives were later invented, allowing for more data storage on smaller devices</li><li>• 2.5-inch hard drives were then invented</li><li>• Currently, solid state drives (using computer chips) are widely used for data storage, not relying on spinning disks.</li></ul>   |

---

In this example, the AI pulled out the key elements of the script and made a simple outline that students can refer to when giving a presentation. In addition to presentations, this can be used in reading courses to create simple outlines of texts. The students can then use the outlines as guides for verbal summary tasks, reinforcing their learning and understanding of the text.

### 3.5. Generating lesson ideas

For veterans and new teachers alike, creating lessons can be challenging. Teachers want to ensure their lessons are educational, meet learning goals, and are engaging. Even teachers who have taught lessons numerous times may find it necessary and challenging to alter them to be more relevant to their learners. LLMs can draw on their vast knowledge of educational fields and aid in the generation of new lesson ideas for a wide range of topics. However, by no means does this

replace the knowledge the teacher holds in educational methodology, pedagogy, or classroom management. Using the AI in this way is similar to brainstorming class ideas with other teachers and determining which would work within their context and aid their learners. Within language learning, teachers also need to consider the language level of their students and develop activities that are accessible and appropriate for their level. In Tables 5 and 6, the AI is prompted with a topic and given a CEFR-level target to adjust the activities for two different classes of students.

Table 5. Example of generating lesson ideas for CEFR B1

|   |
|---|
| <b>Input</b>  |
| You are an EFL teacher. Generate three ideas to act as an introduction to the topic of Media Literacy for CEFR B1 level English learners.   |
| <b>Output (AI response)</b>   |
| Media scavenger hunt: Have students work in pairs or small groups to find examples of different types of media (e.g. advertisements, news articles, social media posts) and analyze them for credibility, bias, and purpose.                            |
| Picture analysis: Show students a series of images and ask them to identify the message being conveyed, who the target audience is, and what techniques are being used to influence the viewer.   |
| Media creation: Have students create their own short piece of media (e.g. a news article, a social media post, a video) and present it to the class for discussion and critique. Focus on how media literacy skills can be applied when creating media. |

Table 6. Example of generating lesson ideas for CEFR C1

|   |
|---|
| <b>Input</b>  |
| You are an EFL teacher. Generate three ideas to act as an introduction to the topic of Media Literacy for CEFR C1 level English learners.   |
| <b>Output (AI response)</b>   |
| Critical media analysis: Choose a current media issue or event and lead a class discussion on the different perspectives and narratives presented in various media outlets. Encourage students to use critical thinking and media literacy skills to evaluate the credibility and reliability of the sources. |

---

Media representation: Have students analyze media representation of a particular group or issue (e.g. gender, race, politics) and discuss the impact of media representation on public perception and attitudes.

Media production: Have students create a media project (e.g. a podcast, a video, a news article) that addresses a current issue or event. Encourage students to consider ethical and responsible media production practices, and to apply media literacy skills in their research and presentation of information.

---

As the examples above demonstrate, the AI is capable of producing multiple lesson ideas for one topic that require varying skill sets. The generated activities range from applying basic topic knowledge to analyzing more complex ideas and creating multimodal projects. Interestingly, the AI is able to scale these lessons depending on the learners' language abilities. In contrast to the CEFR B1 ideas, the CEFR C1 ideas are more cognitively challenging for students. These activity ideas ask that learners have broad knowledge of adjacent topics or the ability to apply knowledge to societal and cultural concepts.

### 3.6. Leveling texts for testing or reading practice

Language teachers commonly have difficulties selecting topically, lexically, and syntactically appropriate texts for their learners. While there are many resources like graded readers or level-appropriate textbooks, it is only sometimes possible to find materials that are motivating for individual learners and appropriately leveled. Teachers often need to source an article from online or a book and modify it for the educational needs of their learners, an extremely challenging and time-consuming process. LLMs are uniquely suited to creating appropriately-leveled topical texts, saving teachers valuable time and effort. By providing the AI with specific instructions about the desired level, length of the passage, and a specific topic, teachers can utilize LLMs to create passages that are well-suited to the needs of the learners (see Table 7).

Table 7. Example of leveling text

---

**Input**

---

In language appropriate for a CEFR B2 English language learner, write a 200-250 word passage about the effect of the TV show CSI on forensic science.

---

---

**Output (AI response)**

---

The TV show CSI (Crime Scene Investigation) has had a significant impact on the field of forensic science. Firstly, it has helped to raise awareness about the importance of forensic evidence in solving crimes and has made people more familiar with the techniques used in forensic investigations. It has also popularized the field, making it more attractive to students who are considering a career in forensic science.

However, the show has also had some negative effects. One of the main criticisms is that it portrays forensic science as a much more glamorous and straightforward process than it really is. This has led to unrealistic expectations among the general public and even among law enforcement officials about what forensic science can and cannot do. The show also often presents forensic evidence in a way that is not scientifically accurate, which can create misunderstandings about the reliability of forensic evidence.

In conclusion, the impact of CSI on forensic science has been both positive and negative. While it has helped to raise awareness and attract new talent to the field, it has also created unrealistic expectations and distorted the public's understanding of the role of forensic evidence in solving crimes. Therefore, it is important for people to understand that the TV show is just a form of entertainment and should not be taken as a representation of the real-life practices of forensic science.

(230 words)

---

In the above example, the AI is able to fulfill the requirements well, meeting the length, language leveling, and topic requirements. Checking the text level using independent tools like *Text Analyzer* (<https://hub.cathoven.com/>) indicates that the text is correctly leveled as CEFR B2. It is important to note that the teacher should verify the generated text's accuracy at this point to avoid the possibility of the AI generating creative, yet untrue information. This issue can be somewhat mitigated by providing clear specific topics, as in the example above. Teachers should also note that the AI does not have knowledge of the most recent news or current events and as such is unable to generate accurate texts on recent topics.

Once teachers have a text, they can also use the conversational memory of ChatGPT to create exercises or comprehension questions based on this text. The example below in Table 8 is based on the text generated in Table 7.

Table 8. Example of question generation for leveled text

---

**Input** (as a follow up in the same conversation thread as above)

---

Generate 4 multiple-choice comprehension questions about the main ideas in the text.

---



---

**Output (AI response)**

---

What is the overall impact of the TV show CSI on forensic science described in the text?

- A. Positive
- B. Negative
- C. Both positive and negative

What is one of the main criticisms of CSI according to the text?

- A. It raises awareness about forensic science
- B. It is scientifically accurate
- C. It portrays forensic science as more straightforward and glamorous than it really is

How does CSI affect the public's understanding of the role of forensic evidence?

- A. It makes people more familiar with forensic techniques
- B. It creates scientifically accurate understanding of the role of forensic evidence
- C. It creates unrealistic expectations

What is the main message of the text regarding CSI?

- A. It is a good representation of real-life practices of forensic science
  - B. It is just a form of entertainment and should not be taken as a representation of real-life practices of forensic science
  - C. It is a negative representation of real-life practices of forensic science.
- 

This example shows that the AI can generate a reasonable set of comprehension questions about a text. Of course, it would be necessary for the teacher to verify the content of the questions before giving them to students. The teacher could further refine the questions by asking the AI to change the number of distractors or to create a question about each paragraph of the text.

#### **4. Discussion and conclusion**

Using an LLM such as ChatGPT to help teachers create activities like the ones we have shared is a clear application of what is possible today with the technology. However, these uses for LLMs are only scratching the surface of what is possible. As educators and learners continue to embrace digital learning, more and more of their output is stored as retrievable data in documents, videos, and voice recordings. There is a growing opportunity for educators to make use of this information to more accurately assess students' abilities and provide them with better guidance. However, the more data we have, the more ways there are to interpret it. No teacher has enough hours in the day, nor days in the week, to sift through each student's treasure trove of data to

make accurate assessments. This is where the true power of AI becomes apparent and relevant to teachers and students as such a task is becoming increasingly trivial for LLMs like OpenAI's GPT. LLMs can save teachers time by analyzing their course outcomes, recommending topics, finding relevant texts, summarizing them, and creating assessment questions for students. Additionally, while students are writing, LLMs can simultaneously analyze grammar, cohesion, and style, while providing feedback and focused study advice. This kind of AI, rather than replacing teachers, has the power to enhance them and provide them with new tools and abilities to help their students focus on areas that need attention.

The activities discussed above demonstrate that LLMs can support teachers by completing tasks on their behalf that could have been overly time-consuming or mundane. Tasks such as correction of grammatical and mechanical errors, as well as adapting texts for learners in level-appropriate language, can be offloaded to LLMs to free up teacher time for other endeavors. Though marking and materials development are a necessary part of teaching in most contexts, they are not only time-consuming but tedious, wasting valuable teacher time that could be spent on personalizing content to meet the needs of learners better. LLMs can easily accomplish these tasks due to their understanding of the conventions of human language, as demonstrated by the activities presented in this article.

LLMs are not limited to mechanical, robotic tasks – and may be used to creative ends. They can also author personalized or unique content, such as generating narrative stories, activity prompts, and adapting texts to new mediums. Though completing these kinds of tasks can be rewarding for the teacher, they are overly time-consuming. Furthermore, while a teacher can create personalized materials for each student, this is simply not practical in the vast majority of language teaching environments. LLMs, fortunately, can accomplish this time-consuming task. This opens up the possibility of creating tailor-made educational experiences, which streamlines the instructional design process to focus on what the students need to progress. Already, AI is employed in educational data mining and intelligent tutoring systems, and as this paper demonstrates, it can be used in practical applications in the language classroom today. LLMs, through their ability to generate natural human-like text and draw from a vast multi-disciplinary knowledgebase, have incredible potential for the field of education, and the activities presented are just the beginning of how educators can use them.

## References

- Brown, A. H., & Green, T. D. (2019). *The essentials of instructional design: Connecting fundamental principles with process and practice*. Routledge.

- Cai, A., Konstantopoulos, D., Davis, R., Zheng, Y., & Liu, D. (2021). Let's code for languages: Integrating AI chatbots into language learning. *The FLTMAG*, November 2021. Retrieved February 19, 2023, from <https://fltmag.com/ai-chatbots/>
- Deng, X., & Yu, Z. (2022). A systematic review of machine-translation-assisted language learning for sustainable education. *Sustainability (Switzerland)*, 14(13). <https://doi.org/10.3390/su14137598>
- Frazier, E., Bonner, E., & Lege, R. (2020). Creating custom AI applications for student-oriented conversations. *The FLTMAG*, November 2020. Retrieved February 19, 2023, from <https://fltmag.com/creating-custom-ai-applications-for-student-oriented-conversations/>
- Fryer, L. K., Ainley, M., Thompson, A., Gibson, A., & Sherlock, Z. (2017). Stimulating and sustaining interest in a language course: An experimental comparison of chatbot and human task partners. *Computers in Human Behavior*, 75, 461–468. <https://doi.org/10.1016/j.chb.2017.05.045>
- Gillani, N. (2023, January 26). ChatGPT isn't the only way to use AI in Education. *Wired*. Retrieved February 19, 2023, from <https://www.wired.com/story/chatgpt-artificial-intelligence-education-networks/>
- Holmes, W., Bialik, M., & Fadel, C. (2019). *Artificial intelligence in education: Promises and implications for teaching and learning*. Center for Curriculum Redesign.
- Kochmar, E., Vu, D. Do, Belfer, R., Gupta, V., Serban, I. V., & Pineau, J. (2022). Automated data-driven generation of personalized pedagogical interventions in intelligent tutoring systems. *International Journal of Artificial Intelligence in Education*, 32(2), 323–349. <https://doi.org/10.1007/s40593-021-00267-x>
- Lahiri, A., Kwatra, V., Frueh, C., Lewis, J., & Bregler, C. (2021). LipSync3D: Data-efficient learning of personalized 3D talking faces from video using pose and lighting normalization. *2021 Conference on Computer Vision and Pattern Recognition*, 2755-2764. Retrieved February 19, 2023, from <https://arxiv.org/abs/2106.04185>
- Lu, X. (2018). Natural language processing and intelligent computer-assisted language learning (ICALL). *The TESOL Encyclopedia of English Language Teaching*, 1–6. <https://doi.org/10.1002/9781118784235.eelt0422>
- Machine learning in education*. (2021). Amazon Web Services, Inc. Retrieved February 19, 2023, from <https://aws.amazon.com/education/ml-in-education/>
- Marche, S. (2022, December 7). The college essay is dead. *The Atlantic*. Retrieved February 19, 2023, from <https://www.theatlantic.com/technology/archive/2022/12/chatgpt-ai-writing-college-student-essays/672371/>
- Moussalli, S., & Cardoso, W. (2020). Intelligent personal assistants: can they understand and be understood by accented L2 learners? *Computer Assisted Language Learning*, 33(8), 865-890. <https://doi.org/10.1080/09588221.2019.1595664>
- Pokrivcakova, S. (2019). Preparing teachers for the application of AI-powered technologies in foreign language education. *Journal of Language and Cultural Education*, 7(3), 135-153. <https://doi.org/10.2478/jolace-2019-0025>
- Roose, K. (2023, January 12). Don't ban ChatGPT in schools. Teach with it. *The New York Times*. Retrieved February 19, 2023, from <https://www.nytimes.com/2023/01/12/technology/chatgpt-schools-teachers.html>

- Settles, B., LaFlair, G.T., Hagiwara, M. (2020). Machine learning–driven language assessment. *Transactions of the Association for Computational Linguistics*, 8, 247-263. [https://doi.org/10.1162/tacl\\_a\\_00310](https://doi.org/10.1162/tacl_a_00310)
- Shukla, S., Shivakumar, A., Vasoya, M., Pei, Y., & Lyon, A. (2019). iLeap: A human-AI teaming based mobile language learning solution for dual language learning in early and special education. *15th International Conference Mobile Learning 2019*, 57-64. <https://files.eric.ed.gov/fulltext/ED601160.pdf>
- Summers, N. (2023). ChatGPT FAQ. *OpenAI Help Center*. <https://help.openai.com/en/articles/6783457-chatgpt-faq>
- Tsai, S. C., Chen, C. H., Shiao, Y. T., Ciou, J. S., & Wu, T. N. (2020). Precision education with statistical learning and deep learning: A case study in Taiwan. *International Journal of Educational Technology in Higher Education*, 17(1). <https://doi.org/10.1186/s41239-020-00186-2>
- Turing, A. (1950). Computing machinery and intelligence. *Mind*, 59(236), 433-460. <https://doi.org/10.1093/mind/LIX.236.433>