

# Generative AI in English Language Education: An Annotated Bibliography

Naomi Takagi

## Introduction

This annotated bibliography explores the use of generative AI tools, particularly ChatGPT, to enhance the quality of English language learning and teaching. Generative AI is a technology that enables the creation of text, images, or other media based on user prompts. ChatGPT, developed by OpenAI, is one of the most prominent examples; it generates text in response to user prompts by drawing on its extensive training data. It can also tailor its output in terms of length, depth, style, and more, according to the user's intentions expressed in the prompt. For example, when prompted to define a word in English, it not only provides definitions but also includes sample sentences to demonstrate its use. When asked to proofread a text, it delivers a revision and highlights key editing points, all within seconds. Additionally, these exchanges can be conducted in various languages. In short, ChatGPT can function like a "tireless language learning assistant" who is multilingual and always at your service (Kohnke, Moorhouse, & Zou, 2023, p. 538).

While generative AI tools like ChatGPT can be a valuable resource for English learners, incorporating it into the language classroom requires thoughtful preparation. First, teachers need to consider how their roles in the course might change. Traditionally, teachers have spent considerable time addressing organizational flow and the correctness of grammar and vocabulary when teaching writing. However, with the introduction of ChatGPT, they may need to adjust their focus, as the tool can now handle these tasks. Moreover, the tool's ability to generate plausible text can easily lead to academic dishonesty, so proper guidance and supervision are necessary to ensure that students' learning is not compromised.

Given the transformative yet controversial nature of ChatGPT and other generative AI tools, a good first step toward successful classroom integration is to learn what researchers, scholars, and teachers have discovered about the tool. Specifically, the current study was launched to seek answers to the following questions:

- What is generative AI, and how are tools like ChatGPT different from other

educational technologies?

- How have students and researchers responded to generative AI?
- What are specific ways in which generative AI is utilized in English language learning and teaching?
- What are the potential dangers of generative AI, and how can teachers guide students to avoid them?

The following lists the materials that have been found useful for answering these questions. They are thematically organized according to the questions above. Some of the materials are directly relevant, while others may seem tangential. Although this study is a modest first step in addressing these questions, I hope it will be helpful for those who are also considering the fruitful integration of generative AI into their classrooms.

### **Annotated Resources**

#### **Context of the Emergence of Generative AI**

Hockly, N. (2023). Artificial intelligence in English language teaching: The good, the bad and the ugly. *RELC Journal*, 54(2), 445-451. <https://doi.org/10.1177/00336882231168504>

In this article, Hockly (2023) discusses a wide range of educational technology (edtech) tools, with generative AI tools being only one aspect. However, this article is valuable for contextualizing them within the broader trend of edtech tools in English language teaching.

During the Covid-19 pandemic, teachers were compelled to rapidly adopt edtech to manage online classes and support students' remote learning. In the post-pandemic era, the reliance on such technology is likely to increase, especially with advancements in artificial intelligence (AI). Today, a variety of AI-based edtech tools are available, including language learning apps like Duolingo, grammar checkers such as Grammarly, and machine translation tools like Google Translate.

While Hockly acknowledges the need for teachers to stay current with these tools, she argues that student privacy and well-being are equally important. She presents examples to highlight these concerns. For instance, some edtech products have reportedly share students' personal data with third-party companies. In another case, a learning management system (LMS) incorrectly flagged a student as struggling when their lack of engagement stemmed from unrelated issues. Furthermore, as mobile technologies become increasingly embedded in students' daily lives, concerns about distraction and disconnection grow, underscoring the need for "attentional literacy" to support students' well-being (p. 450).

Hockly also addresses the inherent risks of using generative AI tools, including ChatGPT. While these tools have been shown to boost students' confidence, motivation,

and self-efficacy, they are prone to biases and irrelevant outputs due to the limitations of the technologies behind them. As a result, Hockly suggests that teachers should carefully supervise students' use of these tools and encourage them to reflect on their experiences. She concludes by emphasizing the importance of exploring best practices for using edtech tools to support language development effectively and ethically.

Although the fast-paced nature of the edtech industry may render some of Hockly's examples less relevant over time, her detailed analysis remains a strong reminder of teachers' responsibility to ensure students' safe use of generative AI and other edtech tools.

Kohnke, L., Moorhouse, B. L., & Zou, D. (2023). ChatGPT for language teaching and learning. *RELC Journal*, 54(2), 537–550. <https://doi.org/10.1177/00336882231162868>

In this article, Kohnke, Moorhouse, and Zou (2023) aim to achieve three main goals: to present the benefits of ChatGPT in language learning, to discuss the ongoing debates and concerns surrounding it, and to advocate for advanced digital literacy to navigate the new reality shaped by this technology.

According to the authors, ChatGPT differs from previous digital assistants like Apple Siri and Amazon Alexa, which perform specific tasks upon request, such as playing the radio or providing weather forecasts. Unlike these one-directional interactions, ChatGPT is designed for continuous, interactive conversations. It employs natural language processing, machine learning, and deep learning to analyze vast amounts of data, provide relevant information, and maintain an ongoing exchange.

Accordingly, ChatGPT can work as though it is a “tireless language-learning assistant,” allowing both teachers and learners to seek help anytime and as often as needed (p. 538). The tasks ChatGPT can perform include the following:

- **Vocabulary expansion:** (1) Defines words in context in both the target and first languages; (2) Provides dictionary definitions, identifies the word's part of speech, and offers sample sentences.
- **Writing assistance:** (1) Generates texts in specific genres, reflecting the designated situation (e.g., email, advertisement, dialogue); (2) Answers questions about grammar and vocabulary.
- **Test/quiz creation:** (1) Generates multiple-choice tests; (2) Creates open-ended test questions.

At the same time, the authors warn of four major issues and drawbacks of ChatGPT:

- **Unethical use (e.g., cheating):** Students may submit ChatGPT-generated work as their own, which could result in inaccurate evaluations.
- **Concerns about originality:** ChatGPT does not provide sources for its responses, raising concerns about plagiarism.

- **Inaccuracy:** ChatGPT may provide incorrect answers while sounding convincing.
- **Bias:** (1) The chatbot’s training data is primarily in English. As a result, when translated into other languages (e.g., Chinese, Japanese), the output may include cultural biases; (2) The training data consists mostly of written language, which may lead to unnatural responses in spoken contexts.

Given ChatGPT’s transformative impact, the authors assert that “advanced digital literacy” is necessary. This new literacy goes beyond basic digital skills, emphasizing the importance of using tools like ChatGPT effectively and ethically while being aware of their limitations. The authors urge educators and institutions to be proactive in guiding students toward a comprehensive understanding of this new technology.

The strength of this article lies in the concrete examples that showcase ChatGPT’s advantages and limitations. The authors present actual outputs, giving readers a clear sense of what the tool can do and where it may fall short. However, some of their points lack specificity. For example, more detail on how the authors reconcile concerns about ChatGPT’s originality—especially regarding its use of others’ work without proper credit—would have been helpful. Additionally, further exploration of the cultural biases ChatGPT might exhibit would have strengthened the discussion. Despite these limitations, this article is a valuable resource for understanding the specific merits and drawbacks of ChatGPT in language learning and teaching.

### **Student and Teacher Perceptions of Generative AI**

Bibi, Z., & Atta, A. (2024). The role of ChatGPT as an AI English writing assistant: A Study of students’ perceptions, experiences, and satisfaction level. *Annals of Human and Social Sciences*, 5(1), 433-443. <https://doi.org/10.35484/ahss>

In this paper, Bibi and Atta (2024) report on research that examines college students’ initial impressions of using ChatGPT in writing. The authors surveyed and interviewed 150 participants and analyzed the datasets thematically to discuss the tool’s impact on students’ confidence, productivity, and writing skills. Their findings indicate that ChatGPT enhances all three areas. The authors highlight that the revision and editing processes present significant challenges to students, as these stages require advanced language expertise, including mastery of syntax, lexicology, and stylistics. They argue that this is one reason writing classes often prioritize technical skills over fostering students’ growth in idea development and communication. By leveraging ChatGPT, writers can bridge these knowledge gaps, gaining greater autonomy in their English writing.

This research sheds light on students’ enthusiasm for ChatGPT through survey data and direct comments. It suggests that students are particularly drawn to using the tool to navigate challenges common in academic writing. However, the research

leaves several questions unanswered. First, the authors do not specify when or how the college students used ChatGPT. Some parts of the discussion imply its use during the revising and editing stages, while others suggest its role in the initial writing process. For example, the authors state, “ChatGPT is an effective tool for generating unique ideas and drafting. . . . Students give various prompts and generate ideas, which improves their first draft’s overall quality. The suggestions lead them to new insights and innovative ideas, and improve their writing” (p. 441). This ambiguity should be addressed, as using the tool for idea generation could lead to a decline in creativity, integrity, and authenticity—key components of academic work. Additionally, the study lacks details about the participants, such as their country of residence. While the authors may assume the generalizability of their findings, providing specific information about the subjects would allow for a more nuanced discussion.

Alm, A., & Ohashi, L. (2024). A worldwide study on language educators’ initial response to ChatGPT. *Technology in Language Teaching & Learning*, 6(1), 1–23. <https://doi.org/10.29140/tltl.v6n1.1141>

Alm and Ohashi (2024) conducted a worldwide study to understand how college-level language educators responded to ChatGPT-3.5 within the first ten weeks following its release in November 2022. With its novel natural language processing capabilities, ChatGPT was expected to impact language education, among numerous other sectors, prompting the researchers to measure its initial effects. A survey collected both quantitative and qualitative data, with responses from 367 educators from 48 countries and regions. The results were analyzed thematically and through the Concerns-Based Adoption Model (CBAM).

Key findings include the following: (1) most respondents were either unaware of or not engaging with ChatGPT, though early adopters had already started using it for educational purposes; (2) nearly 70% of respondents were interested in its educational use, such as creating learning resources, conducting assessments, and facilitating student self-study, but they also expressed concerns about AI potentially replacing the language teaching profession, misuse for plagiarism, and student over-reliance on AI; (3) the CBAM analysis revealed that many educators are in the “Consequence” stage, having already begun considering both ChatGPT’s benefits and potential harms to language learning.

One important contribution of this research is that it measures language educators’ initial responses to ChatGPT through the CBAM framework. This will be valuable as a reference point for future research on evolving stances toward ChatGPT among educators. The survey questions are also insightful, outlining potential applications of ChatGPT in language teaching, such as vocabulary, grammar, writing,

reading, conversation, quizzes, tests, and games. While specific implementation strategies were outside the scope of this research, the findings highlight ChatGPT's potential as a teaching tool.

An unexpected finding is that many respondents were either unaware of or not actively engaged with ChatGPT. The authors attribute this to the research being conducted soon after the tool's release. However, when ChatGPT-3.5 was launched, many educational institutions worldwide reacted swiftly—some with alarm over its potential negative impact on learning, while others embraced it, as seen in Moorhouse's (2024) research. The relatively low engagement in this study may be partly because almost 40% of respondents were based in Japan. It would be interesting for future studies to explore whether there are notable discrepancies in responses across different countries.

Moorhouse, B. L. (2024). Beginning and first-year language teachers' readiness for the generative AI age. *Computers and Education: Artificial Intelligence*, 6, 1-8. <https://doi.org/10.1016/j.caeai.2024.100201>

In this exploratory study, Moorhouse (2024) conducted interviews with pre-service and first-year English language teachers in Hong Kong to explore their perceptions of generative AI (GAI) tools and their readiness to integrate these tools into their teaching practices. The study examined these two groups of teachers separately to gain a fine-grained understanding of their experiences. Additionally, the focus on these groups was driven by the observation that a lack of confidence and competence in technology use is a contributing factor for high resignation rates among new teachers.

The interview data were analyzed in multiple stages, resulting in the identification of five key themes. Below is the breakdown of these themes along with insights shared by the interviewees:

- **GAI Readiness:** Pre-service teachers felt significantly less prepared to integrate GAI tools into their classrooms compared to first-year teachers. One interviewee highlighted the challenge of feeling prepared in an ever-evolving technological landscape, while another noted that the sense of readiness develops only through practical, on-site use of these tools.
- **Awareness of GAI tools and their capabilities and functions:** Since ChatGPT is integrated into a widely used LMS in Hong Kong, first-year teachers were generally familiar with it. However, their understanding varied; some could provide examples of its use and potential pitfalls, while others struggled to grasp its functionality or expressed frustration at their inability to prompt effectively.
- **Utility of GAI tools for language teaching:** First-year teachers reported

using GAI tools for tasks such as designing grammar and vocabulary exercises, creating reading passages, generating teaching ideas, and developing error correction exercises. One teacher expressed amazement at the ability to generate multiple error-correction worksheets within seconds, a task that previously took considerable time.

- **Views towards students' use of GAI tools:** Some schools advised their teachers to require in-class writing to prevent misuse and ensure fairness. Concerns were also raised about students' overreliance on GAI tools, which could diminish their motivation or impede the development of their language competence.
- **Preparing students to use GAI tools productively and critically:** There was a consensus among first-year teachers on the importance of establishing clear policies and guidelines for GAI use. They also emphasized the need to model the productive and critical use of these tools in the classroom.

Based on these findings, the author offers preliminary recommendations. For teachers, it is crucial to engage students in discussions about both the opportunities and challenges presented by GAI tools and to model best practices in their use. For teacher training programs, there is a need to better equip pre-service teachers with the knowledge and skills to effectively integrate GAI tools into their teaching.

Moorhouse's research reveals the importance of access in fostering technological awareness. This is evident in the differing responses from pre-service and first-year teachers regarding their readiness to use GAI tools. Notably, Hong Kong's major LMS integrates ChatGPT, providing a "legitimate and authorized way to use the tool in their professional work" (p. 6). As a result, first-year teachers had greater knowledge and experience with the tool. In this sense, Moorhouse's recommendation for initial teacher education institutions to familiarize pre-service teachers with GAI tools is well-founded. Moreover, educators outside of Hong Kong can also recognize the significance of access in building students' confidence and capability in using technology. Additionally, Moorhouse's findings indicate that some educators may lag behind, even when they have access to technology. Since GAI tools differ from previous technologies, it can take time for some educators to fully benefit from them. This highlights the need for sufficient time and training to support those teachers in their adaptation to new tools.

### **Use of Generative AI in Teaching and Learning**

Tseng, W., & Warschauer, M. (2023). AI-writing tools in education: If you can't beat them, join them. *Journal of China Computer-Assisted Language Learning*, 3(2), 258-262. <https://doi.org/10.1515/jccall-2023-0008>

In this article, Tseng and Warschauer (2023) propose a framework for integrating

AI-based tools into the second language writing classroom. The release of ChatGPT in 2022 put educators world wide on high alert. Its ability to create a text instantly by synthesizing a vast amount of digital data can lead to academic dishonesty like plagiarism. The authors, however, argue that banning them is not the right approach. AI is integrated in various systems and products, rapidly becoming a part of everyday life. Those who are unable to utilize those tools will likely struggle to navigate this new reality. Moreover, since second language learners' late start often fetters their communicative efficacy in the target language, it is imperative that learners know how to leverage the power of AI-based tools. Thus, the authors propose a pedagogical framework, which allows educators to integrate AI literacy education as part of their coursework.

The framework consists of five steps: understand, access, prompt, corroborate, and incorporate. First, students need to understand the capabilities and limitations of AI-based tools. They should also consider whether the use is beneficial for learning and is ethically sound. Second, students need to access those tools and navigate their functions, which is essential to maximizing their benefits. Third, students need to practice framing their prompts as they can determine the kind of answer they receive from those tools. Fourth, they should learn to corroborate. Given the risk of AI-based tools offering incorrect or biased information, students need to critically evaluate the answers by fact-checking when needed. The last element is to incorporate. Students should be held accountable for how they used AI-based tools in their writing process. Lack of guidance may lead to inefficient or unethical use of those tools. By using this framework, Tseng and Warschauer argue that second language educators can enhance language learning, encourage ethical usage, and prepare students for communication beyond classrooms.

While it is evident that students must develop proficiency in AI-based tools to thrive in today's world, teaching their use comes with unique challenges, as it requires a higher level of expertise and responsibility from educators. Tseng and Warschauer's framework, in this regard, offers valuable guidance, pushing us beyond our usual practices. For instance, the first step, "understand," emphasizes the importance of exploring a tool's full capabilities and limitations rather than restricting its use to familiar functions. This stage encourages educators and students alike to be more intentional in their exploration of AI tools. Similarly, the final step, "incorporate," is crucial yet often overlooked, as users may see the tool merely as a means to an end. By including this step, students are prompted to reflect on their use of the tool throughout the learning process, fostering a sense of accountability. The five-step framework presents a systematic and comprehensive approach teachers can turn to as they tackle the challenge of teaching effective and ethical use of AI-based tools.

Bonner, E., Lege, R., & Frazier, E. (2023). Large language model-based artificial intelligence in the language classroom: Practical ideas for teaching. *Teaching English with Technology*, 23(1), 23–41. <https://doi.org/10.56297/bkam1691/wieo1749>

In this article, Bonner, Lege, and Frazier (2023) explore practical ways for language teachers to utilize large language models (LLMs), such as ChatGPT, to create everyday class materials. The authors argue that by leveraging LLMs, teachers can streamline the development of these materials, allowing them to focus more on personalizing instruction and attending to students' individual needs.

To provide context, the authors first discuss the historical background of LLMs, tracing their origins back to the 1950s when British computer scientist Alan Turing envisioned machines with human-like intelligence. The authors highlight key technological advancements, such as machine learning and neural networks, which enabled LLMs to process vast amounts of data, identify patterns, and understand complex relationships. The development of OpenAI's Generative Pre-Trained Transformer (GPT) marked a significant breakthrough, allowing machines to process text in natural, conversational language. ChatGPT, a widely accessible web application with a free version, enables anyone with internet access to generate text efficiently.

The authors suggest that language teachers can leverage LLMs' abilities to synthesize digital data and produce texts based on user prompts. They suggest several practical applications:

- **Summarizing text:** Creating model summaries or simplifying lengthy and complex texts.
- **Grammar and mechanics correction:** Demonstrating corrections for ungrammatical or awkward sentences.
- **Writing prompts:** Providing opening sentences in specific writing modes for students to expand upon.
- **Presentation notes:** Extracting key phrases from presentation scripts.
- **Lesson or assignment ideas:** Generating level-appropriate activities to practice knowledge and skills.
- **Texts for testing and reading practice:** Creating short texts on designated levels and topics.
- **Exercises or comprehension questions:** Formulating questions to assess students' understanding of provided texts.

The authors provide practical tips and cautions for using LLMs. For example, since LLMs' outputs vary based on the specificity of prompts, they recommend that language teachers use detailed instructions, such as specifying word count, CEFR level, and format (e.g., bullet points). They also caution that LLMs may generate confident-sounding responses even when lacking sufficient data. Despite these limitations, the

authors argue that language educators can harness LLMs to enhance both the quality and efficiency of their teaching practices.

Bonner, Lege, and Frazier's discussion offers an in-depth explanation of both the power and challenges of LLMs. For instance, they help us grasp the novelty of LLMs compared to traditional human-computer interactions. Previously, users had to learn and use programming languages such as C, JavaScript, or Python to communicate with computers. Moreover, each language serves different purposes, requiring users to learn multiple languages to accomplish different tasks, which posed a significant barrier for many. With LLMs, however, this dynamic is reversed: computers are now made to understand human languages. As a result, even novices, including many teachers who lack technical expertise, can now communicate with computers to accomplish tasks. Since the tasks introduced above can be time-consuming for teachers, computer assistance could greatly alleviate their burden.

Another key point in the article is the use of the term "large language models" (LLMs) to describe tools like ChatGPT. The word choice is fitting because various LLM tools are already available. More importantly, it underscores both the promise and limitations of these tools. LLMs are trained on vast amounts of language data, and the content is continuously evolving. Consequently, the output varies based on the timing and phrasing of the prompt, highlighting the importance of "prompting" and "corroborating" (Tseng & Warschauer, 2024). While LLMs are increasingly capable of producing desirable outputs, the data they rely on is still human-made, and the system will never be omniscient. The word LLMs itself implies this inherent limitation.

Mizumoto, A., & Eguchi, M. (2023). Exploring the potential of using an AI language model for automated essay scoring. *Research Methods in Applied Linguistics*, 2(2), 1-13. <https://doi.org/10.1016/j.rmal.2023.100050>

Mizumoto and Eguchi (2023) investigated the efficacy of a prompt-based transformer for automated essay scoring (AES). Over the past half-century, AES has garnered significant attention from both researchers and practitioners, as grading essays is not only time-consuming but also prone to inconsistencies due to raters' fatigue, subjectivity, and other factors. Early AES systems relied on superficial text features, such as sentence length and the number of commas. However, advancements in machine learning and deep learning have significantly improved the reliability of AES, to the point that it is now used alongside human rating in high-stake standardized tests like TOEFL and GRE.

Mizumoto and Eguchi, however, argue that the advent of transformers, such as ChatGPT, has the potential to elevate AES to a new level by enabling raters to use AES without any knowledge or skills in computer programming. To test this, they

conducted AES using OpenAI's text-davinci-003 GPT model. They applied the tool to score 12,100 essays from ETS's Corpus of Non-Native Written English, known as TOEFL11, based on the IELTS Task 2 rubric, and compared the results with the benchmark scores. Additionally, the authors incorporated linguistic features that have been widely researched and closely align with human ratings.

The results suggest that prompt-based transformers could be a reliable AES tool. Although the exact matches with the benchmark scores were 54.33%, the consistency rate increased significantly to 89.15% when a small deviation of 1 to 2 points was included. Furthermore, the authors found that when these scores were combined with those based on linguistic features, the consistency rate was even higher. However, while these results are promising, the authors argue that the transformer's imperfections are still evident, underscoring the importance of cross-checking them with human ratings.

Evaluating students' writing is indeed a challenging task for many teachers. As Mizumoto and Eguchi point out, teachers are subject to biases, fatigue, and other influences that make grading difficult and inconsistent. These difficulties are compounded by the fact that the grades they assign affect both the students' overall course grade and their emotional well-being. As a result, teachers may resort to giving uniform grades, except in cases where a submission is significantly below expectations, or they may spend excessive time ensuring their grades are fair. Mizumoto and Eguchi's findings provide hope for teachers, suggesting that machine-generated scores can serve as a valuable reference point, offering an additional layer of objectivity when cross-checking their own assessments.

### **Perils of Relying on Generative AI**

Godwin-Jones, R. (2024). Distributed agency in language learning and teaching through generative AI. *Language Learning & Technology*, 28(2), 5–31. <https://doi.org/10.10125/73570>

In this article, Godwin-Jones (2024) discusses the opportunities and challenges of generative AI tools, particularly ChatGPT. Unlike earlier AI systems that relied on linguistic rules, such as syntax, semantics, and morphology, generative AI uses large language models (LLMs) trained on vast amounts of linguistic data. Through machine learning, these models derive patterns and conventions from statistical data, enabling them to generate coherent responses. ChatGPT is particularly versatile in providing rapid, statistically plausible answers to user input. The latest version, ChatGPT-4.0, released in May 2024, is even more advanced, as it can interpret users' emotions and respond accordingly.

Godwin-Jones argues that, given the transformative impact of generative AI on language learning, educators should teach learners how to effectively use these tools

while developing critical AI literacy. To this end, the author introduces the concept of “shared agency,” where learners recognize that their output often results from interactions with AI rather than their own independent efforts. This understanding can help them recognize the limitations of AI tools as well. For instance, AI lacks pragmatic awareness—understanding the communicative context and the characters and perspectives of the interlocutors. Moreover, relying too heavily on AI early in the writing process can stifle learners’ creativity and critical thinking skills. There are also privacy concerns, as continuous interaction with AI may lead to increased data surveillance. Additionally, AI-generated responses are not always neutral; they tend to reflect biases associated with WEIRD (Western, Educated, Industrialized, Rich, Democratic) perspectives. Godwin-Jones urges educators to understand both the benefits and potential pitfalls of AI tools so they can better prepare learners to navigate an increasingly AI-driven world.

One of the key contributions of this article is the notion of “shared agency.” It helps us recognize that when we interact with generative AI tools, we are engaging in a collaborative process. Moreover, the concept of “shared agency” reminds us that we have control over the extent to which we maintain or relinquish our agency in writing. For instance, in situations where our personal input is less crucial, we may feel more comfortable letting go of some agency. On the other hand, when our presence in the content is vital, relinquishing that control can be detrimental. Writing itself is inherently challenging, as it requires us to make constant decisions about what to say and how to say it. In this sense, using generative AI tools can lead to a slippery slope: while they ease the struggle of writing, increasing reliance on them may risk losing control over our writing, and in educational contexts missing out on valuable opportunities for skill development.

TED. (2024, September 14). What makes us human in the age of AI? A psychologist and a technologist answer [Audio podcast episode]. In *TED Talks Daily*. <https://podcasts.apple.com/jp/podcast/ted-talks-daily/id160904630?i=1000669517059>

In this podcast program, social psychologist Brian S. Lowery and AI technologist Kylan Gibbs discuss what defines our humanity in the age of AI. The conversation begins with Gibbs noting a certain disconnect in interacting with AI. He describes this feeling as missing the spontaneity, genuineness, or contextual understanding — qualities inherent to human communication that make prolonged interactions with AI uncomfortable. Gibbs observes that similar unease occurs when using other new technologies, like virtual reality (VR), social media such as Facebook and Instagram, or even Zoom calls. People want to believe that there is someone real on the other side, but the interactions often feel inauthentic.

Despite these discomforts, the AI user base is growing exponentially. Lowery

suggests this growth might divert us from our fundamental pursuit of human connection. For example, while AI interactions are currently one-on-one, future interactions might involve multiple AIs, potentially isolating people within their individual AI worlds. Additionally, AIs could reinforce users' existing worldviews, leading to rigidity in thought and communication. This could cause people to overlook the broader impact of real human interactions on society and those around them. Lowery and Gibbs conclude by highlighting the potential dangers of AI, suggesting that it could sidetrack us from our essential need to engage and connect with other human beings.

Lowery and Gibbs' discussion provides valuable insights for considering the purpose of language learning in the classroom. With the rise of AI tools, students and teachers may begin to view the goal of language education as mastering the language—correct pronunciation, grammar, and fluency in speaking and writing—assuming that machines can fulfill all their needs. However, without human interaction, such an education risks becoming perfunctory, failing to foster meaningful communication skills. In this sense, Lowery and Gibbs remind us that the true goal of language education goes beyond technical knowledge and skills; it is about cultivating genuine understanding and building connections with others.

NPR. (2024, August 2). Transcript: MIT sociologist Sherry Turkle on the psychological impacts of bot relationships. NPR. <https://www.npr.org/transcripts/g-s1-14793>

One of the applications of generative AI technologies is to assist people with their emotional needs and relationships. Due to its natural language processing and generative capabilities, generative AI is increasingly being used in emotionally intimate contexts such as writing personal letters or creating chatbots that serve as friends. These chatbots can offer unconditional acceptance and affirmation. Some people even create avatars of deceased loved ones to help cope with their grief.

In a podcast interview on *TED Radio Hour*, MIT sociologist Sherry Turkle warns of the hidden danger of relying on AI for emotional support. While chatbots may help alleviate stress and sorrow, they can also deprive individuals of the necessary process of experiencing and working through these emotions. Turkle argues that becoming vulnerable and engaging with difficult emotions are essential to developing and nurturing empathy.

Beyond the impact of generative AI on human intimacy, the interview addresses the broader danger of overreliance on AI in various aspects of life. Although AI chatbots offer convenience and can reduce unwanted stress, Turkle challenges this notion. In response to an anonymous user's claim that dealing with complicated human relationships is too much stress, she responds, "And we need that stress — I guess that's what I'm saying — that stress serves a very important function in our

lives to keep us in our real human bodies and in our real human relationships.”

### Conclusion

This paper has explored the impact of generative AI technology, especially ChatGPT, on language learning. It is remarkable that what once seemed like an impossible dream in the 1950s has come to fruition over half a century later, thanks to key technological advancements in areas such as neural networks, natural language processing, access to massive datasets, and developments in machine learning and deep learning. Today, we can interact with computers in natural language and harness their power to complete tasks.

Students and teachers have embraced this technology, finding creative ways to use these tools. For students, ChatGPT serves as a language tutor on hand, assisting with questions and providing editing support. Teachers have similarly leveraged the technology to streamline tasks like creating worksheets, improving lesson plans, and designing tests, which previously required significant time and effort. While ChatGPT can greatly enhance learning, its use may lead to problems such as intellectual complacency, overreliance, or even plagiarism. Therefore, it is crucial that students receive clear guidance on the ethical use of AI, and teachers must ensure that AI is used to support—rather than replace—the development of essential communication skills.

Although this annotated bibliography has addressed the initial four questions, it only covers a small portion of the available literature. Continued research is essential to examine these questions from different perspectives. Future studies should also focus on methods for integrating AI technology in ways to uphold its primary goal of language learning, which is to help students find the voices to express themselves and develop meaningful connections with others. As with any tool, generative AI’s potential to enhance or hinder learning ultimately depends on how it is used.

### Acknowledgment

The author was assisted by ChatGPT (GPT-4.0) in editing the manuscript to address issues of grammar, phraseology, and style to meet the standards required for scholarly publication. However, the content is entirely the author’s original work, and the author is solely responsible for any errors present in this paper.

### References

- Alm, A., & Ohashi, L. (2024). A worldwide study on language educators’ initial response to ChatGPT. *Technology in Language Teaching & Learning*, 6(1), 1–23. <https://doi.org/10.29140/tltl.v6n1.1141>
- Bibi, Z., & Atta, A. (2024). The role of ChatGPT as an AI English writing assistant: A Study

- of students' perceptions, experiences, and satisfaction level. *Annals of Human and Social Sciences*, 5(1), 433-443. <https://doi.org/10.35484/ahss>
- Bonner, E., Lege, R., & Frazier, E. (2023). Large language model-based artificial intelligence in the language classroom: Practical ideas for teaching. *Teaching English with Technology*, 23(1), 23-41. <https://doi.org/10.56297/bkam1691/wieo1749>
- Godwin-Jones, R. (2024). Distributed agency in language learning and teaching through generative AI. *Language Learning & Technology*, 28(2), 5-31. <https://doi.org/10.10125/73570>
- Hockly, N. (2023). Artificial intelligence in English language teaching: The good, the bad and the ugly. *RELC Journal*, 54(2), 445-451. <https://doi.org/10.1177/00336882231168504>
- Kohnke, L., Moorhouse, B. L., & Zou, D. (2023). ChatGPT for language teaching and learning. *RELC Journal*, 54(2), 537-550. <https://doi.org/10.1177/00336882231162868>
- Mizumoto, A., & Eguchi, M. (2023). Exploring the potential of using an AI language model for automated essay scoring. *Research Methods in Applied Linguistics*, 2(2), 1-13. <https://doi.org/10.1016/j.rmal.2023.100050>
- Moorhouse, B. L. (2024). Beginning and first-year language teachers' readiness for the generative AI age. *Computers and Education: Artificial Intelligence*, 6, 1-8. <https://doi.org/10.1016/j.caeai.2024.100201>
- NPR. (2024, August 2). Transcript: MIT sociologist Sherry Turkle on the psychological impacts of bot relationships. *NPR*. <https://www.npr.org/transcripts/g-s1-14793>
- TED. (2024, September 14). What makes us human in the age of AI? A psychologist and a technologist answer [Audio podcast episode]. In *TED Talks Daily*. <https://podcasts.apple.com/jp/podcast/ted-talks-daily/id160904630?i=1000669517059>
- Tseng, W., & Warschauer, M. (2023). AI-writing tools in education: if you can't beat them, join them. *Journal of China Computer-Assisted Language Learning*, 3(2), 258-262. <https://doi.org/10.1515/jccall-2023-0008>

## Generative AI in English Language Education: An Annotated Bibliography

高木 有美

### 要旨

本研究は、生成AI、特にChatGPTの英語教育現場への導入の可能性を探ることを目的としている。生成AI技術を使った代表的なツールであるChatGPTは、2022年にリリースされて以来、利用者が急速に増加し、その影響は日本の大学教育にも波及している。同時に、その使用方法に対する懸念も広がっており、本研究は、このツールを教育現場に効果的に導入する方法を検討するものである。具体的なテーマは以下の通りである。(1) 生成AIとは何か、そしてChatGPTのようなツールは他のデジタルツールとどのように異なるのか、(2) 生成AIに関して学習者や研究者はどのように考えているか、(3) 生成AIは英語学習および指導において具体的にどのように利用されているのか、(4) 生成AIの潜在的な危険性とは何か、そしてそれらをどのように回避すべきか。

これらの質問に答えるために選定された資料は、主に学術論文であるが、生成AIの進展に批判的な視点を提供するポッドキャスト番組も含まれている。本研究で取り上げた資料からは、生成AIが数十年にわたるコンピュータ科学の進歩を集約した画期的な技術であり、その活用次第では、生徒の自信、生産性、自主性を高めることに大きく寄与する可能性が示唆された。しかし同時に、過度な依存、学術的不正、アルゴリズムの不完全性、さらには学習者の主体性喪失や感情面への影響も指摘されている。また、これらの課題に対する主な対策として、デジタルリテラシーの強化や生成AIツールの現場導入における枠組の構築が提案されている。本研究は、レビューされた資料の数は限られているものの、生成AI技術を英語教育の現場に取り入れるための準備の一助となることが期待される。