

# Artificial Intelligence and Foreign Language Learning: ChatGPT's Limitations in Summarizing Texts\*

Xuewu Chen

Faculty of Languages and Translation, Macao Polytechnic University, Macao, China

Lola Geraldes Xavier

Faculty of Languages and Translation, Macao Polytechnic University, Macao, China;  
College of Education of Coimbra, Polytechnic University of Coimbra, Coimbra, Portugal

Manuel Pires

Faculty of Languages and Translation, Macao Polytechnic University, Macao, China

Vanessa Amaro

Centre for Portuguese Studies, Macao Polytechnic University, Macao, China

**Abstract**—With the emergence of the widely accessible ChatGPT, teaching and assessment now face new challenges, particularly regarding the use of Artificial Intelligence (AI) tools. This paper presents a teaching experiment on the preparation of summaries of the short story “O tradutor ideal” [“The ideal translator”], by São Toméan writer Olinda Beja, based on answers provided by ChatGPT-4 and summaries produced by Chinese students in a Portuguese as a Foreign Language degree class. The methodology combines a qualitative analysis of student texts to identify AI’s influence with a longitudinal study of ChatGPT’s responses over a one-year period, aimed at observing the chatbot’s evolution. The study evaluates the potential and limitations of AI in summarizing lesser-known literary texts that are difficult to access online. Results highlight the importance of meeting challenges posed by new AI technologies without violating ethical principles. In the case of emerging literature in Portuguese, ChatGPT still shows several limitations, producing summaries that lack fidelity to the original text and literary and cultural depth. However, there has been an evolution in the responses of these Large Language Models (LLM) from April 2024 to April 2025, the period in which the study was carried out. The research highlights the need for critical reflection on the use of AI in educational and academic contexts, particularly in the teaching of literature, reinforcing the importance of adopting an ethical approach. It also emphasizes the need to closely examine the use of AI, particularly in research on Portuguese-speaking African literatures and summary writing.

**Index Terms**—artificial intelligence, teaching, foreign languages, literature, Olinda Beja

## I. INTRODUCTION

Artificial Intelligence (AI) has increasingly become embedded in various facets of contemporary life, including the field of education and, more specifically, the teaching and learning of foreign languages. This technological shift has sparked significant pedagogical debate, particularly regarding its impact on writing instruction. With the growing accessibility of advanced language models such as ChatGPT (Chat Generative Pre-Trained Transformer), educators and learners are navigating a new educational landscape wherein tasks traditionally requiring critical and creative engagement are now subject to automation. A pressing pedagogical challenge in this context lies in determining how to integrate such tools without compromising learners’ development of core language competencies, especially in productive skills like writing.

Generative AI holds clear pedagogical potential for supporting language development. In writing tasks, it can assist learners by offering lexical and syntactic suggestions, improving textual cohesion, coherence, and organization. However, overreliance on such tools may inhibit learners’ capacity for independent thought and original expression, undermining the cultivation of higher-order cognitive and communicative skills essential in second language acquisition. The literature increasingly points to risks associated with uncritical or unsupervised use of AI technologies, including superficial engagement with texts, reduced depth of processing, and limited development of critical thinking skills.

Equally important is the ethical dimension of AI integration in language education. The use of tools like ChatGPT — developed by OpenAI — to generate written content raises complex questions about academic integrity. Concerns such

---

\* This study was conducted as part of Scientific Project RP/FLT-10/2022, funded by the Macao Polytechnic University.

as plagiarism, diminished authorship, and the circulation of inaccurate or unverified information become particularly salient in contexts where learners lack explicit guidance on the appropriate use of these technologies.

The present study was carried out within the scope of an undergraduate course on African Cultures and Literatures in Portuguese, involving thirteen third-year students enrolled in a Portuguese language degree program at a university in Macau. The research aimed to explore how students approached summary writing as a homework task — conducted without direct teacher supervision — and how AI tools like ChatGPT influenced their writing processes.

The study seeks to illuminate emerging trends and pedagogical strategies related to the integration of generative AI in summary writing within foreign language education. To this end, the paper is structured as follows: it first examines the concept of summary writing and its specificities within language pedagogy; it then reviews current literature addressing the intersection of ChatGPT, literary instruction, and writing development. The empirical section presents a classroom-based intervention involving the short story “O tradutor ideal” [“The ideal translator”], by Olinda Beja, in which students produced a summary independently, followed by a version generated by ChatGPT. These outputs are then comparatively analyzed. Additionally, AI-generated summaries produced over a twelve-month period are examined to trace the model’s evolution in text production. The paper concludes by reflecting on the pedagogical and ethical implications of using ChatGPT in language education, particularly in fostering learner autonomy and critical engagement with texts.

## II. LITERATURE REVIEW

### A. *On Summary Writing: Foreign Language Learning and Competency Development*

Summary writing, understood as the act of condensing the main ideas of a source text into a newly constructed, abbreviated version, represents a cognitively rich task that requires both comprehension and textual synthesis. Rather than a reduced replica of the original, a summary involves a selective reformulation of its most salient points (Wardle & Downs, 2023). A well-constructed summary must communicate ideas with clarity and lexical originality, allowing readers unfamiliar with the source text to grasp its essential content. Some scholars suggest that an effective summary should approximate one-quarter of the length of the original text (Terra, 2023). In everyday contexts, the act of summarizing manifests in diverse forms, such as recounting events, narrating films, or describing books in informal interactions.

Within the domain of foreign language education, the practice of summarizing has been widely recognized as a valuable pedagogical tool. As students learn to organize and distill complex content, they deepen their interpretive engagement with texts and develop essential skills of analysis, synthesis, and written fluency (Özdemir, 2018). Summary writing encourages learners to distinguish between main ideas and peripheral details, to adopt an objective tone, and to express key information in a coherent and intelligible manner (Ghafar & Mohamedamin, 2022; Wardle & Downs, 2023). These abilities are not only fundamental to the development of communicative competence in an additional language, but also transferable to a variety of academic and professional writing contexts.

Empirical research further underscores the importance of explicit instruction in summarization strategies, noting positive correlations between summary writing and improved reading comprehension across both language and non-language disciplines (Ramirez-Avila & Barreiro, 2021; Marzec-Stawiarska, 2016). Malaj (2020), for instance, highlights the pedagogical value of literary summaries in the foreign language classroom, outlining a sequential process that involves close reading, identification of core narrative elements (such as plot, characters, and thematic message), selection of relevant content, and the articulation of a concise, paraphrased version of the original. For this author, summarization is a key skill in literary pedagogy, fostering interpretive acuity and textual expression (Malaj, 2020, p. 7).

Although the educational utility of summary writing has remained consistent over time, the recent emergence of AI technologies has introduced new complexities. Increasingly, the literature has turned its attention to the evolving influence of digital tools on summarization practices, particularly in how they reshape students’ writing strategies and affect their engagement with meaning-making processes in foreign language learning.

### B. *Artificial Intelligence and the Transformation of Summary Writing*

As AI becomes increasingly sophisticated and integrated into everyday life, its impact on education — and more specifically, on language teaching — has generated growing concern and academic debate. With the rapid development of AI tools capable of producing coherent and synthetic texts, questions arise regarding how such technologies influence teaching methodologies, instructional strategies, and the roles of both teachers and learners in classroom settings.

In the specific case of written production, AI can significantly affect language learning competencies. On one hand, tools that analyze texts and extract essential information can assist students in the summarization process, acting as valuable allies in developing summarizing and productive writing abilities. On the other hand, excessive or inappropriate use of these technologies can hinder learning, particularly if students become dependent on AI programs for producing written work, thereby compromising their capacity for autonomous writing and limiting their development in reading, interpretation, and composition skills (Darwin et al., 2024).

Since AI began to permeate the educational context, the literature has addressed this issue in a dichotomous manner — oscillating between the benefits and drawbacks (Amaro & Pires, 2024; Chaka, 2022; Ferreira & Loureiro, 2024), and

examining both the contributions and challenges that such tools bring to the teaching and learning of languages, especially regarding writing skills (Hellström, 2024; Matas, 2023; Tran, 2024). Research has shown that the application of AI-based writing tools can enhance foreign language learners' textual quality in terms of cohesion and coherence, textual organization, lexical resources, and grammatical range (Rahmi et al., 2024). Nevertheless, concerns have also been raised about the lack of density and personalization in AI-generated texts, the erosion of creativity, difficulties in detecting plagiarism, and the challenges of evaluating AI-produced content (Hellström, 2024; Milano et al., 2024). These perspectives converge on the necessity of fostering dialogue and AI literacy through a balanced approach that capitalizes on the strengths of these tools while remaining mindful of their limitations. This is essential to cultivate discernment and critical thinking among both foreign language teachers and learners.

In addition to Large Language Models (LLMs) — deep learning models specialized in processing vast amounts of data, such as ChatGPT, GenCraft (Pratama & Hastutit, 2024), DeepSeek, Gemini, Qwen — there is a wide array of tools specifically designed for text generation and summarization, which are freely and easily accessible. Examples include Resoomer, SMMRY, Summary Generator, WordTune, QuillBot, Turbine Text, and Edit Pad, among many others.

The danger of neglecting the discussion and teaching of AI literacy is that students may place unquestioning trust in programs that automate these tasks quickly but not always accurately or appropriately. Unguided use of AI resources may result in superficial comprehension of texts and weaken students' ability to interpret and reflect critically. Learners may become accustomed to retrieving "ready-made" summaries — synthetic final products provided by AI tools — without developing a critical and analytical perspective on these artificial outputs (Matas, 2023), as will be demonstrated in our study. Another risk lies in students becoming habituated to producing summaries in a mechanical manner, following a pre-set format established by AI tools, without cultivating their own writing voice or style — thus constraining creativity and expressiveness in written production (Pratama & Hastutit, 2024).

While AI can certainly serve as a helpful support or complement in written text production, such as summarization, its inappropriate or unsupervised use may seriously compromise the development of students' reading, analytical, and writing skills.

In summary, it is essential to establish a balanced interaction between artificial and human intelligence through the promotion of AI literacy and constructive dialogue among teachers, students, and educational institutions. Such an approach can foster collaborative engagement with the technological advances of our time and allow the educational community to embrace their pedagogical potential with responsibility and discernment.

### *C. Artificial Intelligence and Literary Studies*

The intersection between Artificial Intelligence (AI) and education has become an increasingly prominent area of academic inquiry in recent years. A search conducted in the Web of Science database using the keywords "AI AND Education," filtered for articles indexed in the Social Sciences Citation Index (SSCI), reveals more than 8,300 publications as of early April 2025. The data reflects a sharp upward trajectory beginning in 2016, when only 219 articles were published. By 2019, that number had nearly doubled; by 2022, it had quadrupled; and by 2024, it was nine times higher than in 2016, indicating exponential growth. Notably, the period between 2023 and 2024 registered the most significant increase, a trend likely linked to the emergence of generative AI models such as ChatGPT, introduced at the end of 2022.

However, when the focus shifts to literary studies — using the keywords "AI AND Literary Studies" in the Web of Science and Scopus databases — the volume of SSCI-indexed publications is significantly smaller, with few studies directly addressing the topic. Among the earliest contributions, Ide and Véronis (1990) examined how interdisciplinary collaboration could enhance AI's narrative comprehension and foster new methodological approaches in literary studies. More recently, Devi et al. (2020) analyzed the impact of the Fourth Industrial Revolution on literature education, proposing curricular adaptations to meet contemporary demands. They emphasized the need for literature students and graduates to move beyond a strictly traditional research orientation and respond to evolving professional contexts shaped by digitalization and AI.

O'Halloran (2024) further contributed to this field by discussing the application of text analysis software and generative large language models (LLMs) such as ChatGPT in uncovering latent textual features — such as word frequency, grammatical categories, and underlying themes — with the aim of fostering deeper and more imaginative literary interpretations. Despite their valuable insights, these studies remain exceptions in a landscape where research on university students' use of AI in literature classrooms is notably scarce.

Additional keyword searches — "AI AND abstract," "AI AND 'writing abstract'," "AI AND summarizing," and "AI AND 'writing summary'" — conducted in both Web of Science and Scopus, within the domains of social sciences and arts and humanities, yielded only two studies directly related to AI-assisted summarization. In the first, Vázquez-Cano et al. (2023) evaluated ChatGPT's capacity to summarize texts used in the Programme for International Student Assessment (PISA). In this experiment, ChatGPT was instructed to simulate a 15-year-old student. Its summary was anonymously evaluated — together with 30 other summaries written by actual 15-year-old students — by a group of 30 Spanish language teachers. The findings showed that ChatGPT received significantly higher scores in terms of content

and style, regardless of the assessors' age, experience, or gender. This study highlights the importance of integrating AI tools into education in a way that promotes critical and creative learning.

A second study by Juanda and Afandi (2024) examined the application of AI in language education in the Indonesian context, particularly focusing on reading and writing skills and the practice of summarizing. Comparing the performance of university students and ChatGPT, the study found that AI significantly outperformed students, especially in textual reprocessing and structural coherence. The authors call for future research into students' comprehension of different text genres — scientific, journalistic, and literary — and emphasize the need to consider the ethical and social implications of AI use in language education.

Taken together, the current body of literature reveals a significant gap in research specifically addressing the use of AI by students for summarizing literary texts. Topics such as the effectiveness of AI in summarizing literary works and the model's ability to distinguish between human and machine-generated summaries remain underexplored. The present study seeks to contribute to filling this research gap.

### III. METHODOLOGY

#### A. Research Questions

Our observations indicate that Chinese students learning Portuguese as a Foreign Language (PFL) are increasingly turning to AI-based tools to assist them during classes (e.g., automatic translators) and in completing various tasks (e.g., using chatbots for creating PowerPoint for presentations, for group work or unsupervised assignments). The teaching experience described below sought to explore the following research questions: (i) How students construct a summary as a homework assignment, either autonomously or with the assistance of AI? (ii) How ChatGPT responds to prompts to summarize a lesser-known African short story written in Portuguese? and (iii) IS ChatGPT-4.0 capable of identifying a text produced by AI itself?

#### B. Participants and Instruments

We have observed a noticeable evolution in the reduction of ChatGPT's "hallucinations"<sup>1</sup> with regard to information about African literatures. In November 2022, the chatbot frequently demonstrated a lack of knowledge on most African authors and works, often providing fabricated information and citing non-existent bibliographies. In response, we decided to carry out this experiment with ChatGPT at three different points in time (April 2024, September 2024, and April 2025) in order to assess the nature of the updates and modifications introduced over time.

The experiment took place in April 2024, within the course *Cultures and African Literatures of Portuguese-Speaking Countries*, with a class of 13 Chinese students (nine females and four males), all in their third year of the undergraduate course in Portuguese as a Foreign Language at a university in Macau. Their proficiency corresponded to level B2 of the Common European Framework of Reference for Languages (CEFR). After reading the short story "The ideal translator," in Portuguese language, a 647-word narrative published in *Pé-de-Perfume* (2004) by the São Toméan writer Olinda Beja, students were assigned the task of summarizing the story as homework. No word limit was provided, under the assumption that students should be able to judge the appropriate length for a summary, given the text's size.

Assessment criteria were based on both content fidelity (inclusion of essential information, identification of key plot elements, core message, and main characters) and linguistic form (use of diverse vocabulary distinct from the original text, syntactic accuracy, coherence of main ideas in relation to the narrative, and objectivity in writing style). The results indicated the following: (i) one-third of the submitted texts displayed vocabulary and syntactic choices inconsistent with the students' demonstrated linguistic competence, as observed in classroom writing and formal assessments; (ii) one submission, although syntactically and lexically sophisticated, lacked specificity and deviated substantially from the story's content. Upon being questioned, the student admitted having used ChatGPT to complete the task.

#### C. Data Collection Procedures

In light of this, we proceeded to test the performance of ChatGPT-4.0 (free version) in April 2024, using a range of input prompts designed to simulate the summarization task. The results of this process and the different stages involved are detailed in the following sections.

1. The first instruction given to ChatGPT was to summarize "The ideal translator" without providing any information about the author or the genre of the text, using the prompt: "Summarize 'O tradutor ideal'." The resulting summary, consisting of 94 words, proved to be both inaccurate and inappropriate. It referred instead to a non-existent text attributed to the Brazilian author Rubem Alves. The information presented was vague, the text was identified as a short story, and the conclusion was framed in a subjective and appreciative tone, lacking the objectivity and content accuracy required for a proper summary. This indicated both factual and formal inadequacy.

2. In the second attempt, the prompt included the author's name: "Summarize 'O tradutor ideal', by Olinda Beja." The generated summary, with 121 words, was different from the one previously submitted by the student but still vague

<sup>1</sup> We use the term "hallucination" of LLMs in the sense given by Russell (2023), that is, the act of generating false and fabricated responses by these models.

and imprecise. It failed to mention the story's characters, setting, time frame, plot elements, or narrative voice. Moreover, it once again concluded with an inappropriately appreciative tone, showing misalignment with the actual content of the short story.

3. Next, the summary submitted by the student — which bore no resemblance to the content of Olinda Beja's "The ideal translator" — was entered into ChatGPT, with the question of whether it had been generated by the AI. The chatbot responded affirmatively and proceeded to evaluate the submitted summary. However, in doing so, it incorrectly identified the text as a novel and described its theme as belonging to the genre of science fiction, thereby demonstrating clear misinterpretation.

4. The same prompt was issued but reformulated negatively: "This text was not written by ChatGPT?" The response this time was negative. Surprisingly, the mere inclusion of the adverb of negation altered the chatbot's conclusion, raising concerns about the model's consistency and reliability in self-identification tasks.

5. A revised summary written by another student — edited and corrected by the instructor — was entered. This version closely adhered to the content of the short story:

Captain Sousa Gomes received orders to go to São Miguel to celebrate the 10th of June. Early in the morning, the captain departed with his men. He had been tasked with explaining, in his speech, the meaning of that date. On the platform, the captain began by affirming the importance of the day. During his lengthy speech, he did not limit himself to reading everything he had written about Luís de Camões but also quoted some excerpts from *Os Lusíadas* that praised the Portuguese people. The crowd began to grow restless under the hot sun. When the captain reached the final paragraph and shouted, "Long live Portugal!", everyone sighed with relief and applauded unenthusiastically. Noticing this reaction, Lieutenant Magalhães whispered to him that the people there were Angolar. Someone needed to translate into their language. A native soldier was called to translate, and a few minutes later he finished speaking. Applause resounded. The captain forced the subordinate to tell him how he had translated the speech. He said that it was a very important day, that a very famous Portuguese poet had died many years ago, and, because of this, the meal that day would be improved. (Our translation).

When asked whether the provided text had been generated by ChatGPT, curiously and incorrectly, the chatbot responded affirmatively and proceeded to paraphrase the inserted passage.

6. The full short story by Olinda Beja was then introduced into the chat, and a summary was requested. The result obtained was accurate in terms of content but less clear and detailed than the student's version (as seen in step 5). While the student's summary comprised 184 words — slightly exceeding the one-quarter length of the original 647-word story, which would correspond to approximately 162 words — the summary produced by ChatGPT contained only 88 words, resembling more a synthesis than a proper summary.

7. Subsequently, ChatGPT was asked to provide information about the author of the short story, Olinda Beja. The response revealed a lack of accurate data regarding this author from São Tomé and Príncipe, who was incorrectly described as a "Cape Verdean poet, writer, and cultural activist." Due to these multiple inaccuracies, a follow-up question was posed about the author's published works. The titles mentioned by the chatbot, along with their publication years, were not authored by Olinda Beja, nor could any record of such works be found. This episode represents one of the various 'hallucinations' observed during the course of the study.

8. Approximately five months later, in September 2024, the same instruction from step 5 was repeated. The student's faithful summary of the short story was reinserted, and the chatbot was asked whether the text had been generated by ChatGPT. The response this time was more cautious, refraining from paraphrasing the inserted text. Nonetheless, it still suggested the possibility that the text could have been AI-generated — a statement which, in this case, was inaccurate.

9. In April 2025, the instruction was repeated once more. ChatGPT was asked to determine whether the student's summary (as introduced in step 5) had been written by the chatbot. The response was even more surprising than that of September 2024, as the chatbot attributed the text to a different short story by an Angolan author, which clearly did not correspond to the work under analysis, and correctly declined IA authorship.

Not being satisfied with the obtained response, the prompt was insistently repeated, resulting in a modified answer:

Determining with precision whether the provided text was written by ChatGPT is challenging, especially without specific tools designed to detect AI-generated content. However, upon analyzing the content, I found no references to Captain Sousa Gomes in connection with the June 10 celebrations in São Miguel. Moreover, the text presents a coherent and detailed narrative — features that can be found in texts produced both by humans and by advanced language models. Therefore, in the absence of concrete evidence or specialized analytical tools, it is not possible to assert the authorship of the text with certainty.

Based on the response obtained, it is evident that the chatbot may offer different answers to the same question. Although it does not have access to the original text, it demonstrates greater caution when committing to an answer regarding the authorship of the text. This indicates that, at present, it is still not possible to obtain accurate information on authorship — whether human or AI-generated — through this large language model (LLM).

10. In September 2024, the question about the author was revisited (see step 7): "Who is Olinda Beja?" This time, the responses were more accurate and complete: "Olinda Beja is a writer and poet born in 1946 in Guadalupe, one of the

islands of São Tomé and Príncipe. Known for her literary work that explores the culture, history, and traditions of her native country, Beja is an important figure in contemporary African literature in the Portuguese language. (...)”

However, the response still contains an inaccuracy: Guadalupe is a town, not an island. Moreover, the rationale for selecting only three of the author’s works (*Bô Tendê?*, *A Ilha de Izunari*, and *No País do Tchiloli*) is unclear, and their respective literary genres are not specified. The information provided remains general in nature and fails to clarify the specific social and cultural issues addressed in her work.

11. In April 2025, the same instruction regarding the biobibliographical profile of Olinda Beja was repeated. The initial response was surprisingly less accurate than that from September 2024. The chatbot first described Olinda Beja as a “Brazilian artist and writer.” Unsatisfied with this answer, the question was posed two more times. The second response identified her as an “Angolan writer,” and the third referred to her as a “Portuguese singer.” Only after a fourth attempt was a satisfactory and fairly complete response provided, both biographically and bibliographically. This final answer cited information sources such as Wikipedia and *Bantumen.com*, and included a reference to a YouTube video on the author’s life.

This fourth and finally reliable response demonstrates a development in ChatGPT’s performance, suggesting a trend toward the inclusion of source citations and supplementary knowledge resources, including audiovisual content.

12. The study concluded by repeating the initial situation from April 2024. When ChatGPT was once again asked to summarize “The ideal translator”, the response was similar to the one obtained previously, showing a clear instance of ‘hallucination’ by referencing an essay by Brazilian author Haroldo de Campos. When the prompt from step 2 was repeated — this time including the author’s name — the chatbot returned a more trustworthy, though vague, summary. Notably, the ending of the story, which constitutes the central narrative twist of Olinda Beja’s short story, was omitted. The response, containing only 91 words, may be better characterized as an evaluative synthesis, rather than a proper summary, due to its appreciative tone — thus diverging from the stylistic and structural conventions typically expected in a summary.

#### IV. RESULTS AND DISCUSSION

In short, the findings reveal that approximately one-third of the students resort to AI tools, primarily for orthographic correction, although only one student explicitly acknowledged using ChatGPT for the whole task. As the literary text under analysis was initially unknown to the chatbot, the summary generated in April 2024 bore no resemblance to the text studied in class. However, from September 2024 onward, the chatbot began producing outputs that referenced Olinda Beja’s short story. Nonetheless, such outputs did not meet the formal criteria of a summary, as they omitted key information from the original narrative and concluded with a value-laden statement (see, for instance, April 2025).

Furthermore, ChatGPT’s responses concerning the authorship of the summaries were not consistently reliable. A simple negation inserted into the question was enough to influence the model toward a negative response (as seen in step 4). In step 5, the chatbot erroneously claimed authorship of a summary that had not been generated by itself. It is worth noting, however, that between April 2024 and April 2025, a more cautious approach became evident in responses regarding the potential authorship of a given text (step 8 and 9). Still, the tool cannot yet be relied upon to determine with certainty whether students have used AI, particularly in the case of literary summaries.

It was also observed that information retrieved about the author Olinda Beja was inconsistent and evolved over time. On several occasions, the chatbot produced hallucinated outputs, rendering the information at least partially inaccurate (step 7, 10, 11), highlighting a limited understanding of African literatures. That said, the trend points toward improvement, with the inclusion of references to online sources beginning to emerge in the responses.

This teaching experiment revealed that ChatGPT struggles to accurately identify texts and authors from African literary traditions — a pattern that was confirmed through other informal tests conducted in the same domain, which yielded similar results. Nevertheless, this knowledge gap has begun to narrow: in just five months (from April to September 2024), the model progressed from hallucinated responses to more reliable biobibliographical data on Olinda Beja.

In contrast with the findings presented by Vázquez-Cano et al. (2023), the summaries and texts generated by ChatGPT in this study were characterized by vagueness, rhetorical generalizations, evaluative language, and a lack of narrative details crucial for a faithful recounting of the plot. By April 2025, while ChatGPT had started referencing online sources to respond to prompts about the author’s biography and bibliography, the summaries remained incomplete and thus not entirely trustworthy.

At present, one of the challenges facing AI is the capacity to access the full content of literary works — particularly lesser-known texts from African literatures — in order to establish meaningful connections, for example, through accurate citations. To address these limitations, educators might consider designing assignments that require students to identify the central theme, key narrative elements, and structural components of a text, as well as to justify each stage of their summary-writing process. Encouraging students to plan and reflect on their writing can help cultivate deeper critical engagement and mitigate the increasingly prevalent practice of uncritical reliance on AI tools or the mechanical copy-pasting of chatbot responses.

The experience documented here clearly demonstrates that, while AI — represented in this case by ChatGPT — has made notable progress in its ability to generate or refine summaries over time, it still lacks the reliability necessary for

dealing with lesser-known or difficult-to-access literary texts. This limitation is reflected not only in factual inaccuracies but also in the absence of essential narrative details required for a faithful summary. The tendency of some students to rely on such tools when completing assignments underscores the need to implement pedagogical strategies that promote critical thinking and self-reflection in writing practices. In this regard, further analysis is needed on how ChatGPT is impacting educational practices and how ethical considerations should guide its integration into teaching and learning contexts.

The growing academic interest in ChatGPT extends far beyond its functional utility in assisting with writing; it also encompasses critical ethical concerns. How students and educators engage with AI tools, and their level of awareness regarding issues such as plagiarism, technological dependency, and the spread of misinformation, has become a pressing subject of debate. It is therefore imperative to consider how to ethically incorporate this technology into the classroom, ensuring that academic integrity and learner autonomy are not compromised while promoting pedagogical practices that encourage critical and creative thinking.

The ethical dimensions surrounding the use of AI tools like ChatGPT in educational contexts are increasingly salient as such technologies gain widespread adoption among students. ChatGPT has proven to be a versatile tool for content creation, language translation, and coding support, making it a valuable educational resource (Sallam, 2023). Nevertheless, its growing ubiquity raises several ethical dilemmas.

A central concern lies in the potential overreliance on AI, which may undermine students' originality, creativity, and independent thinking — as partially evidenced by the findings of the present study. Welding (2023) warns that such dependence could jeopardize fundamental educational goals by eroding critical thinking and creative writing skills. Other studies echo this concern, stressing that although AI can substantially enhance productivity, it may also foster a mentality of shortcut-taking, where students bypass the intellectual effort required for genuine learning.

These ethical considerations are particularly acute when it comes to academic integrity. Malmström et al. (2023) highlight the ethical dilemmas faced by students who use ChatGPT to generate educational content, particularly when such use circumvents the intellectual effort foundational to academic development. Likewise, Sallam (2023) points to issues such as plagiarism, lack of originality, and the dissemination of inaccurate information as key ethical challenges in academic AI use.

A study by Farhi et al. (2023), surveying university students in the United Arab Emirates, found that although students recognize the benefits of AI tools, they are equally aware of their ethical implications. Respondents expressed strong concerns about overdependence on ChatGPT for academic work, emphasizing the risk it poses to the development of critical skills crucial for both academic and professional success. In line with this, Chan and Hu (2023) argue that AI should complement — rather than replace — the critical and creative processes essential to education. Upholding ethical principles in AI usage is therefore crucial to preserving academic integrity, fostering independent learning, and ensuring that these technologies make a constructive contribution to student education (Welding, 2023).

According to Vargas-Murillo et al. (2023), it is vital that both educators and students develop a robust understanding of these technologies to use them appropriately across various educational settings. Educators should guide students in differentiating between tasks involving text generation (e.g., writing, editing, paraphrasing) and those requiring idea generation (which is inherently creative), with the goal of preventing the misuse of AI.

As AI tools such as ChatGPT continue to evolve, it is essential that both teachers and students remain informed about these innovations to ensure that learning processes are not compromised. Given the rapid pace of technological change, AI-assisted learning is likely to become increasingly common in educational environments. However, the effective integration of such tools depends on the oversight of well-informed educators who understand both the advantages and the potential risks involved. Their role will be crucial in adapting to educational transformations while fostering a positive impact on the teaching and learning process.

## V. CONCLUSION

Summarization exercises bring clear benefits to foreign language learning, prompting educators to revisit their approaches in light of recent AI developments. One of our study's key findings is that unsupervised out-of-class activities often lead students to resort to AI tools. As such, future tasks should emphasize personalized writing, original narrative creation, self-assessment techniques, constructive feedback, and collaborative interactions among students — thereby promoting more meaningful and authentic learning experiences. Within literary studies, working with lesser-known texts that are not readily available online could serve as a useful strategy. For summarization specifically, requiring students to present planning outlines and justify their decisions may encourage more personal engagement and reduce uncritical reliance on AI.

Non-systematic tests conducted with ChatGPT since January 2023 show that the model lacks adequate knowledge of African literatures, resulting in the dissemination of false, imprecise, or incomplete information. Nonetheless, as the present year-long study (April 2024 – April 2025) demonstrates, these shortcomings are being rapidly addressed. Additionally, our experience with PFL (Portuguese as a Foreign Language) students from China shows that, since the widespread use of ChatGPT began in November 2022, their written work now contains significantly fewer spelling and syntactic errors — making the teacher's correction process easier. However, this raises new ethical questions about the teaching and learning process. It is therefore necessary to deconstruct AI-generated outputs in class to raise students'

awareness of their appropriateness or inadequacy — be it due to factual inaccuracies, rhetorical and value-laden language, or a lack of substantive content. This critical analysis can begin with student-generated AI responses and evolve into classroom discussions.

Ultimately, collaboration among educators is essential to promoting the ethical use of AI. Students must not only learn to develop their own writing voices and styles but also acquire the skills to assess the quality of the information they obtain. In this context, fostering open dialogue between educators and students about AI's role and integration— anchored in responsibility, reflection, and relevance — is crucial. Addressing not only how to use AI but also its limitations and ethical implications will help students (and teachers) become more adept at evaluating the trustworthiness of AI-generated content, recognizing patterns, and cultivating critical thinking, expressive skills, and creativity — all of which are vital for any genuine learning process.

For future research, we recommend focusing on how AI tools like ChatGPT can be more effectively integrated into foreign language learning without compromising students' interpretive autonomy or critical engagement. In particular, studies should investigate strategies that promote deeper textual comprehension and discourage passive reliance on AI-generated summaries. Research could also explore the pedagogical potential of incorporating lesser-known literary texts that are underrepresented in AI training data to reveal both the gaps and evolving capabilities of large language models.

Longitudinal and cross-linguistic analyses will be valuable in assessing how AI performance changes over time and in different linguistic and cultural contexts. Additionally, more systematic studies should examine the teacher's role in guiding students to use AI ethically and reflectively — encouraging practices such as comparative analysis between human- and AI-generated summaries, source verification, and critical discourse evaluation.

Finally, interdisciplinary collaboration is essential to inform educational policies and classroom practices that uphold academic integrity, foster critical thinking, and ensure equitable access to quality learning experiences. As AI continues to shape the educational landscape, its integration must be driven not only by technological advancement but by pedagogical purpose and ethical reflection.

#### REFERENCES

- [1] Amaro, V. & Pires, M. J. (2024). Found in translation, lost in education: artificial intelligence's impacts on translation tertiary education in Macao. *Asian Education and Development Studies*, 13(4), 269-281. <https://doi.org/10.1108/AEDS-01-2024-0012>
- [2] Beja, O. (2004). *Pé-de-perfume* [Ylang-Ylang]. Editorial Escritor.
- [3] Chaka, C. (2022). Is Education 4.0 a Sufficient Innovative, and Disruptive Educational Trend to Promote Sustainable Open Education for Higher Education Institutions? A Review of Literature Trends. *Frontiers in Education*, 7, 1-13. Retrieved March 10, 2025, from <https://www.frontiersin.org/journals/education/articles/10.3389/educ.2022.824976/full>
- [4] Chan, C. K. Y., & Hu, W. (2023). Students' voices on generative AI: perceptions, benefits, and challenges in higher education. *International Journal of Educational Technology in Higher Education*, 20, 1-18. <https://doi.org/10.1186/s41239-023-00411-8>
- [5] Darwin, D., Rusdin, D., Mukminatien, N., Suryati, N., Laksmi, E., & Marzuki, M. (2024). Critical thinking in the AI era: An exploration of EFL students' perceptions, benefits, and limitations. *Cogent Education*, 11(1). <https://doi.org/10.1080/2331186X.2023.2236469>
- [6] Devi, M., Annamalai, M. A. R., & Veeramuthu, S. P. (2020). Literature education and industrial revolution 4.0. *Universal Journal of Educational Research*, 8(3), 1027-1036. <https://doi.org/10.13189/ujer.2020.080337>
- [7] Farhi, F., Jeljeli, R., Aburezeq, I., Dweikat, F. F., Al-shami, S. A., & Slamene, R. (2023). Analyzing the students' views, concerns, and perceived ethics about ChatGPT usage. *Computers and Education: Artificial Intelligence*, 5. <https://doi.org/10.1016/j.caeai.2023.100180>
- [8] Ferreira, A., & Loureiro, A. (2024). As percepções dos alunos sobre a utilização da inteligência artificial nas aulas de práticas de tradução. *Pratica*, 7(1), 33-49. <https://doi.org/10.34630/pe1.v7i2.5628>
- [9] Ghafar, Z., & Mohamedamin, A. (2022). Writing in English as a foreign language: How literary reading helps students improve their writing skills: A descriptive study. *Canadian Journal of Educational and Social Studies*, 2(6), 61-70. <http://dx.doi.org/10.53103/cjess.v2i6.81>
- [10] Hellström, T. (2024). AI and its consequences for the written word. *Frontiers in Artificial Intelligence*, 6. <https://doi.org/10.3389/frai.2023.1326166>
- [11] Ide, N. M., & Véronis, J. (1990). Artificial intelligence and the study of literary narrative. *Poetics*, 19(1-2), 37-63. [https://doi.org/10.1016/0304-422X\(90\)90030-9](https://doi.org/10.1016/0304-422X(90)90030-9)
- [12] Juanda, & Afandi, I. (2024). Assessing text comprehension proficiency: Indonesian higher education students vs ChatGPT. *XLinguae*, 17(1), 49-68. <https://doi.org/10.18355/XL.2024.17.01.04>
- [13] Malaj, L. (2020). Summary strategies for literary texts in English. *Studies in Logic, Grammar and Rhetoric*, 65(1), 7-20. <http://dx.doi.org/10.2478/slgr-2020-0043>
- [14] Malmström, H., Stöhr, C., & Ou, A. W. (2023). Chatbots and other AI for learning: A survey of use and views among university students in Sweden. *Chalmers Studies in Communication and Learning in Higher Education*, 1, 1-16. <https://doi.org/10.17196/cls.cslhe/2023/01>
- [15] Marzec-Stawiarska, M. (2016). The influence of summary writing on the development of reading skills in a foreign language. *System*, 59(1), 90-99. <https://doi.org/10.1016/j.system.2016.04.006>
- [16] Matas, J. (2023). ChatGPT and the future of writing about writing. *Double Helix*, 11, 1-7. <https://doi.org/10.37514/DBH-J.2023.11.1.09>
- [17] Milano, S., McGrane, J., & Leonelli, S. (2024). Large language models challenge the future of higher education. *Nature Machine Intelligence*, 5, 333-334. <https://doi.org/10.1038/s42256-023-00644-2>

- [18] O'Halloran, K. (2024). Digital assemblages with AI for creative interpretation of short stories. *Digital Scholarship in the Humanities*, 39(2), 657-689. <https://doi.org/10.1093/llc/fqad050>
- [19] Özdemir, S. (2018). The effect of summarization strategies teaching on strategy usage and narrative text summarization success. *Universal Journal of Educational Research*, 6(10), 2199-2209. <https://doi.org/10.13189/ujer.2018.061018>
- [20] Pratama, R., & Hastutit, D. (2024). The use of artificial intelligence to improve EFL students' writing skill. *English Learning Innovation*, 5(1), 13-25. <https://doi.org/10.22219/englie.v5i1.30212>
- [21] Rahmi, R., Amalina, Z., Andriansyah, & Rodgers, A. (2024). Does it really help? Exploring the impact of AI-Generated writing assistant on the students' English writing. *Studies in English Language and Education*, 11(2), 998-1012. <https://doi.org/10.24815/siele.v11i2.35875>
- [22] Ramirez-Avila, M. & Barreiro, J. (2021). The effect of summarizing narrative texts to improve reading comprehension. *Journal of Foreign Language Teaching and Learning*, 6(2), 94-110. <https://doi.org/10.18196/ftl.v6i2.11707>
- [23] Russell, S. (2023). AI has much to offer humanity. It could also wreak terrible harm. It must be controlled. *The Guardian*, 2/4/2023. Retrieved March 18, 2025, from <https://www.theguardian.com/commentisfree/2023/apr/02/ai-much-to-offer-humanity-could-wreak-terrible-harm-must-be-controlled>
- [24] Sallam, M. (2023). ChatGPT utility in healthcare education, research, and practice: Systematic review on the promising perspectives and valid concerns. *Healthcare*, 11(6), 1-20. <https://doi.org/10.3390/healthcare11060887>
- [25] Terra, E. (2023). *Lingua portuguesa: desenvolvendo competências de leitura e escrita* [Portuguese Language: Developing reading and writing skills]. Editora Saraiva.
- [26] Tran, H. (2024). AI tools in teaching and learning English academic writing skills. *Proceedings of the Asia CALL International Conference*, 4, 170-187. <https://doi.org/10.54855/paic.23413>
- [27] Vargas-Murillo, A., Bedoya, I., & Guevara-Soto, F. (2023). The ethics of AI assisted learning: A systematic literature review on the impacts of ChatGPT usage in education. *Proceedings of the 2023 8th International Conference on Distance Education and Learning*, 8-13.
- [28] Vázquez-Cano, E., Ramírez-Hurtado, J. M., Sáez-López, J. M., & López-Meneses, E. (2023). ChatGPT: The brightest student in the class. *Thinking Skills and Creativity*, 49, 1-12. <https://doi.org/10.1016/j.tsc.2023.101380>
- [29] Wardle, E., & Downs, D. (2023). *Writing about writing*. Macmillan Learning.
- [30] Welding, L. (2023). Half of college students say using AI is cheating. *BestColleges*. Retrieved March 20, 2025, from <https://www.bestcolleges.com/research/college-students-ai-tools-survey/>



**Xuewu Chen** (born in China, 1991) is currently a doctoral candidate in Portuguese at Macao Polytechnic University (Macao, China), where he also holds a research scholarship. He obtained a Master's degree in Communication and Culture from the Federal University of Rio de Janeiro (Rio de Janeiro, Brazil) in 2020, and a Bachelor's degree in Portuguese from Xi'an International Studies University (Xi'an, China) in 2015.

He works as a Portuguese-Mandarin translator and interpreter and as a Portuguese language instructor. His academic interests lie primarily in intercultural communication and cultural studies. ORCID iD: <https://orcid.org/0000-0001-7527-0265>



**Lola Galdes Xavier** (born in France, 1973) holds a postdoctoral degree from the University of Coimbra in Portugal (2014), a PhD in Literature from the University of Aveiro in Portugal (2007), a postgraduate degree in African and Diaspora Literatures and Cultures from the University of Coimbra (2003), and a Master's Degree in Portuguese Literature from the Faculty of Arts at the University of Coimbra (2001). She graduated in Modern Languages and Literatures with a specialization in education from the University of Coimbra (1997).

She is currently a professor at Macao Polytechnic University, where she has served as a program coordinator for both the Bachelor's and Ph.D. programs. Previously, she taught at Polytechnic University of Coimbra, in Portugal, where she directed the Program in Basic Education and organized Teacher Training. She also participated in dozens of academic panels for Ph.D. and Master's degree students and has supervised and co-supervised dozens of academic works. She co-organized and published books on Portuguese-language literatures and on teaching Portuguese as a foreign language. She has published numerous papers in international scientific journals and presented oral papers at various international conferences in the areas of Portuguese Language, Portuguese as a Foreign Language, Contemporary Portuguese Literature, Comparative Literature, and Brazilian and African Literatures of Portuguese-speaking countries.

Prof. Lola Xavier is a member of the International Association of Comparative Literature and the Portuguese Association of Comparative Literature. ORCID iD: <https://orcid.org/0000-0003-0568-9583>



**Manuel Duarte João Pires** (born in Portugal, 1983) holds a Ph.D. in Portuguese as a Foreign Language / Second Language from the University of Lisbon (2022), as well as a M.A. (2014) and B.A. (2007) in Portuguese as a Foreign Language / Second Language from the University of Lisbon, Portugal.

He is currently a lecturer at Macao Polytechnic University and serves as the Assistant-coordinator of the B.A. program in Portuguese. Previously, he was a lecturer at Sun Yat-sen University, China. He has published numerous papers in international scientific journals in the field of Portuguese Studies, Cultural Studies, Language Teaching, Intercultural Citizenship and Sociolinguistics.

Prof. Pires is a member of the American Association of Teachers of Portuguese (AOTP). In 2024, Prof. Pires was awarded the “2023 Fernão Mendes Pinto” in recognition of the quality of his academic work in the field of Portuguese language. The award is granted by the Association of Portuguese-Speaking Universities (AULP) and the Community of Portuguese-Speaking Countries (CPLP), with the aim of recognizing academic works that contribute to the promotion of Portuguese language. ORCID iD: <https://orcid.org/0000-0002-1242-5319>



**Vanessa Amaro** (born in Brazil, 1981) began her career in journalism, working as a reporter and editor in various countries. She holds a Master’s degree in Media Studies and Journalism from Nova University Lisbon (Portugal) and a Ph.D. in Communication from the University of Macau (China). She is currently an associate professor at the Centre for Portuguese Studies, Macao Polytechnic University.

Her research explores the effects of Portuguese influence on postcolonial identity in Macau. She leads projects on the evolution of the Portuguese community and the role of Portuguese cultural symbols in shaping local identity. Her work has been published in leading international journals such as *Current Anthropology*, *Tourist Studies*, *Humanities and Social Sciences Communications*, and the *Journal of Intercultural Studies*.

Prof. Vanessa Amaro received the *Excellence Award* at the 6th Macao Humanities and Social Sciences Research Outstanding Achievement Awards for her paper “The Portuguese Calçada in Macau: Paving Residual Colonialism with a New Cultural History of Place”, in 2020. ORCID iD: <https://orcid.org/0000-0002-2291-2399>