The use of ChatGPT as a learning tool to improve foreign language writing in a multilingual and multicultural classroom

Stavros Athanassopoulos1 Polyxeni Manoli2 Maria Gouvi3 Konstantinos Lavidas4 Vassilis Komis2
1 Department of Philosophy, University of Patras, Patras, Greece
2 Department of Educational Sciences and Early Childhood Education, University of Patras, Patras, Greece
3 School of Humanities, Hellenic Open University, Patras, Greece

Abstract: The association of Artificial Intelligence (AI) and education has attracted researchers’ attention, as it is very beneficial for learners affecting the whole teaching procedure. ChatGPT, an AI language model developed by OpenAI and based on the Generative Pre-trained Transformers (GPT)-3.5 architecture, though newly introduced in November 2022, became very popular even within the first two months. Researchers have already explored and acknowledged the affordances of ChatGPT for language learning and teaching. In this context, the study aimed to examine the effectiveness of ChatGPT as a feedback tool on foreign language writing, especially for socially vulnerable populations, such as refugees/migrants, who face extra language difficulties. The study was conducted in a Junior High School (Gymnasium) classroom in an urban city in south Greece, which consisted of 8 students with a migrant/refugee background aged 15 years old who were attending German as a foreign language learning. The focus of the study was to analyze how the use of ChatGPT could improve foreign language writing in terms of vocabulary and grammar. After the students’ first writing attempt, we asked ChatGPT to enhance the vocabulary and grammar of the texts. In contrast, the students received an improved version in terms of language and grammar, and then, they were asked to produce another writing piece of work on a similar topic to their first attempt. The study results indicated an increase in the total number of words, the unique words and the average word number per sentence in the improved version of their work. Overall, the study results were promising, indicating that ChatGPT could act as a language learning tool and support the teaching procedure, especially for students with a refugee/migrant background. However, further research is needed to investigate its use and effectiveness in diverse contexts.

Keywords: artificial intelligence, ChatGPT, foreign language learning, writing skills, language education for refugees/migrants

1 Introduction

Artificial intelligence (AI), according to McCarthy’s definition (McCarthy, 2007), is often referred to as the “father of AI”, which is the science and engineering of making intelligent machines brilliant computer programs (Karakose et al., 2023). It is related to the similar task of using computers to understand human intelligence, but AI does not have to confine itself to biologically observable methods (p.2). The beginning of the conversation on AI was marked in 1950 by Alan Turing’s seminal work “Computing Machinery and Intelligence”, by setting the question, “Can machines think?” (Turing, 1950). Since then, he has been offering a test known as the “Turing Test”, a method for determining whether a computer has intelligence. Depending on the tasks an AI system is designed and trained to complete, it can be categorized as either Weak AI or Strong AI (What is Strong AI? IBM, 2023). The first category includes systems that can perform a specific task, such as virtual personal assistants and Industrial robots.

At the same time, the Strong AI indicates that a machine has intelligence comparable to a human being with self-aware consciousness and can solve problems, learn, and plan for the future. Considering that humans are the most intelligent creatures, AI aims to create systems that simulate human acts intelligently (What Is Artificial Intelligence and How Does AI Work? 2023). For every human function, several computer science technologies attempt to imitate that act intelligently (TechEmergent, 2023). For instance, humans’ ability to write and read texts in a language is the field of Natural Language Processing (NLP). Additionally, humans can listen and speak, which belongs to speech recognition; because much of speech recognition is statistically based, it is referred to as statistical learning. In this frame, various projects, as the augmentor (Lavidas et al., 2023) apply AI that takes advantage of big data and learning analytics to give...
multiple stakeholders intelligible proposals for effectively finding and recognizing educational resources, as well as for creating customized learning profiles that take into consideration individual actors’ qualities, needs, and inclinations (Karakose et al., 2023).

Moreover, the human brain can acquire knowledge, generate new ideas, and make decisions based on the inside and machines, known as machine learning. At the same time, neural networks are the tool for their implementation. Depending on the nature of the data and the desired outcome, machine learning can be categorized as supervised, unsupervised, semi-supervised or reinforced learning. As mentioned above, NLP is the AI technology that simulates human writing and reading. Sentiment Analysis, Machine Translation, Speech Recognition, Chatbots and Virtual Assistants are some applications that use this technology. In particular, Chatbots are software applications that perform specific functions, imitating human written and spoken speech as well as possible. Namely, software answers our questions automatically and is used in various contexts, such as customer support, e-commerce, and personal assistance. Their functionality ranges from rule-based chatbots to more advanced AI-based chatbots.

Very recently, on November 30, 2022, a new chatbot called ChatGPT appeared and became trendy quickly. ChatGPT has over 100 million users within the first two months (NapSaga, 2023). Although chatbots have been used in educational contexts since the early 1970s, ChatGPT’s potential to simulate human conversation as realistically as possible opens a new and exciting path in language learning (Huang et al., 2022; Kim et al., 2022; Kohnke et al., 2023). ChatGPT was developed by OpenAI, drawing on various cutting-edge technologies to implement an advanced language model that uses deep learning and NLP techniques to generate humanlike texts in excellent quality according to the input provided. More specifically, it is an improved version that builds upon the successes and lessons learned from earlier versions, such as GPT-3, which have employed the Transformer architecture and have yielded remarkable results in various NLP tasks, such as text generation, translation, summarization, problem-solving, content and grammar correction and question answering (OpenAI, 2023). Its ability to work on complicated tasks that require human intelligence, such as answering exam-style questions in business, law, and medicine or writing academic texts, has placed it as an innovative learning tool, attracting researchers’ interest (Zhai, 2022).

Moreover, ChatGPT supports multiple languages, including English and Korean and can modify the content and format of the generated text to meet the user’s demands and needs, simultaneously creating a natural conversational experience. The most striking feature of ChatGPT is that it can generate text based on the user’s instruction, which means that the desired answer can be gradually modified and developed depending on the users’ request for the reader to be more detailed or succinct (Kim et al., 2023). In this context, the affordances of ChatGPT for language learning and teaching have already been explored and acknowledged by researchers, indicating promising results.

2 Literature review

Since the study focuses on exploring the use of ChatGPT in education, particularly L2 learning, the following literature review will be mainly restricted to it. However, many researchers have explored the potential applications of ChatGPT in various fields, such as customer service, mental health, social media, etc., highlighting its potential to revolutionize multiple areas and improve human-AI interactions in diverse settings.

2.1 Education

Education aims to obtain knowledge; however, comparing education of the past with education of the future, this aim has been modified considering the advancement of technology and the digital era offering education a wide range of semiotic resources and communication channels focusing on how to search for information and combine and edit knowledge (Cope & Kalantzis, 2009). The use of ChatGPT, the latest offspring of technology, in education has been investigated by researchers, indicating a wide range of capabilities. In particular, in mathematics, Pardos and Bhandari (2023) explored ChatGPT’s effectiveness in producing hints compared to those made by human tutors in algebra. They revealed that both human tutors and ChatGPT had positive learning outcomes, highlighting that only the human tutor-generated hints resulted in statistically significant gains. Simultaneously, Frieder et al. (2023) examined its potential in mathematical problems, indicating that ChatGPT’s math skills are much lower than those of a typical graduate student in mathematics. As an AI language model, broader language uses of ChatGPT have been explored. For instance, Feng et al. (2023) studied ChatGPT as a sentence simplification tool and found that the results were close to the simplified sentences produced by humans.
Moreover, Jiao et al. (2023) investigated the ability of ChatGPT to translate. They indicated that it performed similarly to commercial translation products, such as Google Translate, for more widespread European languages but struggled with less widespread ones. They suggested using the source sentence into a more widespread pivot language before translating it into the target language. Mitrovic et al. (2023) probed into the way ChatGPT could answer different questions by using specific patterns and found that it tended to describe experiences rather than express feelings, use uncommon (unusual) words while refraining from using personal pronouns, aggressive and rude language, which are usually attributed to human-generated texts. Furthermore, Zhai (2022) investigated whether the ChatGPT could write an academic paper, revealing that ChatGPT could assist researchers in academic writing while reducing the time to complete the whole writing process within 2-3 hours. Another line of research (Khaliil & Er, 2023) investigated whether essays produced by ChatGPT could be detected for plagiarism through the use of two popular plagiarism detection tools and showed that ChatGPT written productions avoided detection by plagiarism detection software while having a high potential for producing sophisticated text outputs.

2.2 ChatGPT and language learning

Researchers have already explored and acknowledged the affordances of ChatGPT for language learning and teaching (Jurayev, 2023; Sukmayasa & Sudiana, 2023). However, the number of relevant studies still needs to be improved, as ChatGPT has been launched recently. To begin with, Kohnke et al. (2023) discussed its affordances and indicated some uses in language learning by providing examples of learning tasks produced by ChatGPT, such as text production of various genres, text translation, grammar and content correction, vocabulary explanation in multiple ways, production of reading comprehension questions (De La Cruz et al., 2023), which could be used by teachers and learners arguing that AI-driven digital tools, like ChatGPT, are becoming increasingly important in language education. Other studies, for example, Kim et al. (2023) discussed the functions and misconceptions of ChatGPT, supporting that teachers and educational institutions should be encouraged to change their mentality to adapt teaching and assessment in consideration of the inevitable progress of technology. Kim et al. (2023) explored its effectiveness in designing course content based on Task-Based Language Teaching (TBLT). They indicated promising results, suggesting that ChatGPT has the potential to act as a language-learning tool.

Moreover, Baskara and Mukarto (2023) discussed its use in language courses to generate texts on various topics. They indicated that ChatGPT’s potential for language learning is significant, such as providing personalized language instruction and developing authentic language material for learners, highlighting, simultaneously, concerns about the ethical implications of the substitution of human language teachers and the use of a machine learning system to generate text. Similarly, Yan (2023) examined the use of ChatGPT in an L2 writing practicum, finding that it had the potential to improve efficiency in writing and contribute to L2 writing pedagogy, expressing, at the same time, concerns about threats to academic honesty and educational equity. Hong (2023) explored the potential benefits of ChatGPT for foreign language teachers and the issues and risks associated with its use, arguing that ChatGPT offers significant opportunities for teachers and educational institutions to improve language teaching and assessment (Sholekah et al., 2022). Despite some drawbacks, mainly focusing on ethical implications, ChatGPT has been seen as a versatile and valuable tool with significant potential to support learners and promote engaging and adaptive language learning. At the same time, simultaneously, researchers encourage language teachers to explore its pedagogical affordances.

3 Methods

3.1 Research rationale aims and questions

Considering, on the one hand, the contribution of AI technology to language learning and, on the other hand, the first promising findings of the use of ChatGPT as an L2 learning tool (e.g., Hong, 2023; Kim et al., 2023; Kohnke et al., 2023; Yan, 2023), the present study aimed to explore the effectiveness of ChatGPT as a feedback tool on L2 writing, especially for refugees/migrants, who face extra language difficulties. According to relevant research, refugees/migrants who belong to socially vulnerable populations differ not only in age, gender, and socio-cultural background but also in first language(s), language proficiency, and pre-settlement experiences, making teachers’ work even more challenging (Beacco et al., 2017; Dryden-Peterson, 2015; Kantzou et al., 2017). This socially vulnerable population often needs extra help to overcome language barriers and learn an L2, whose inclusion in the educational system of a hosting country can be challenging. At the same time, the number of pertinent
studies on its use in language learning remains relatively limited, as Chat GPT has been recently launched. Therefore, there is a need for further research to unravel ChatGPT’s potential and limitations in language learning. In this context, the following research question was formulated:

To what extent can the use of ChatGPT contribute to improving L2 writing in terms of vocabulary and grammar for students with a refugee/migrant background?

3.2 Research participants and educational setting

The study was conducted in a Junior High School (Gymnasium) in an urban city, Patras, in south Greece. Two hundred forty-seven students attended the specific school, while its staff consisted of 35 teachers. The school included reception classes for 42 students with special educational needs, while 50 students with a refugee/migrant background attended the mainstream classes. The school director and most educators have worked in this academic setting for a long time (from 5 to 22 years) and have gained experience working with students from diverse cultural and linguistic backgrounds.

More specifically, 8 Junior High School students aged 15 with a refugee/migrant background participated in the study. Their countries of origin varied, as most came from Albania, while others came from Italy, Russia and the Philippines. Seven of the participants were females, while one student was male. Four participants were born and raised in Greece, while the rest have lived in Greece for ages ranging from 8 to 12 years. All participants could speak Greek proficiently. The research participants attended the third grade of Junior High School, particularly the German class, which was offered twice weekly (45 minutes for each category). They have been learning German as a foreign language for five years (level A2 according to the sixth level scale of the CEFR for Languages). The course book used in the third grade of High School was “Deutsch – ein Hit 2” “Issue A”.

3.3 Research procedure

The research procedure took place for two weeks and was thematically integrated into the currently used syllabus, fitting with the main principles of the Integrated Foreign Languages Curriculum (IFLC, 2016). More specifically, at the first stage of the research, the students were asked to write a personal email to a German friend describing their favourite sport. Then, the ChatGPT was used to correct their written outputs regarding grammar and vocabulary. ChatGPT was particularly asked to improve their written texts’ vocabulary and grammar. In the second stage, the students received an improved version of their original emails regarding language and grammar. They were asked to study them thoroughly while discussing all the revisions and suggestions offered by ChatGPT, which students should pay attention to improve their final work. At the final stage, students were asked to be involved in the writing procedure by addressing a similar topic to their first attempt; namely, they were asked to write a personal email to a German friend expressing their thoughts about sports and healthy life. The choice of a related topic aimed to motivate students to use new but pertinent words and expressions used by ChatGPT in the improved texts to foster their writing skills.

4 Results

For a more objective comparison of each student’s texts, we measured several qualitative characteristics, including Total Words, Unique Words, Average Words Per Sentence, and Most Frequent Words. These measurements were performed using the same ChatGPT platform to which we provided the text and requested the corresponding analyses. Subsequently, for each qualitative characteristic, we calculated the mean of the optimization across all students, as well as the best and worst-case scenarios.

According to the results of the research, it was revealed that after using ChatGPT as a feedback tool in L2 writing, there was an increase in the total number of words used by students in the improved version of their work. In particular, the increase in the total number of words ranged from 2% in the case of Student 1 to 62% in the case of Student 6, while the mean of the increase was 38%. Simultaneously, it was found that there was an increase in the unique words that students used in their improved version of their work, with an average mean of 19.1 per student, except for Student 1, whose final text had one unique word less than his first draft. On the contrary, the maximum increase was observed in the case of student 6, who had 36 more unique words than in his initial work. The findings also showed that students made longer sentences using more words, which indicated improved L2 writing skills. More specifically, the mean average of words per sentence was approximately 42%, while only Student 4 had a decrease. Finally, in the last column of Table 1, we indicate each student’s most frequently used words in the pre- and post-intervention texts.
### Table 1 Results before and after ChatGPT use

<table>
<thead>
<tr>
<th>Student</th>
<th>Total words Before</th>
<th>Total words After</th>
<th>Unique words Before</th>
<th>Unique words After</th>
<th>Average Words Per Sentence Before</th>
<th>Average Words Per Sentence After</th>
<th>Most frequent words</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>90</td>
<td>92</td>
<td>63</td>
<td>62</td>
<td>5.3</td>
<td>9.2</td>
<td>Mannschaft (3)</td>
</tr>
<tr>
<td>2</td>
<td>91</td>
<td>127</td>
<td>75</td>
<td>89</td>
<td>6.1</td>
<td>11.5</td>
<td>Volleyball (2)</td>
</tr>
<tr>
<td>3</td>
<td>113</td>
<td>146</td>
<td>97</td>
<td>89</td>
<td>10.3</td>
<td>9.7</td>
<td>Volleyball (3)</td>
</tr>
<tr>
<td>4</td>
<td>100</td>
<td>160</td>
<td>99</td>
<td>99</td>
<td>8.3</td>
<td>9.4</td>
<td>Lieblingssport (2)</td>
</tr>
<tr>
<td>5</td>
<td>84</td>
<td>136</td>
<td>89</td>
<td>89</td>
<td>7.0</td>
<td>12.4</td>
<td>Lieblingssport (3)</td>
</tr>
<tr>
<td>6</td>
<td>91</td>
<td>122</td>
<td>84</td>
<td>84</td>
<td>6.1</td>
<td>8.7</td>
<td>Volleyball (2)</td>
</tr>
</tbody>
</table>

5 Discussion

ChatGPT, an AI language model, is regarded as a multifaceted learning tool, as it can support language learning in various ways, such as text generation in multiple genres (e.g., stories, recipes, emails), translation, summarization, problem-solving, content and grammar correction, vocabulary enhancement and question answering (OpenAI, 2023) for many languages, skills, topics, and levels (e.g., Huang et al., 2022; Kim et al., 2022; Kohnke et al., 2023). Thus, ChatGPT provides learning experiences in various contexts depending on learners’ needs (Kim et al., 2023).

Considering the first promising findings of the use of ChatGPT as a language learning tool (e.g., Hong, 2023; Kim et al., 2023; Kohnke et al., 2023; Yan, 2023), the aim of the present study, which was a pilot study, was to explore how the use of ChatGPT could improve foreign language writing in terms of vocabulary and grammar in a multilingual and multicultural classroom. In particular, after the students’ first writing attempt, the researchers asked ChatGPT to improve the vocabulary and grammar of the texts, while the students received an improved version in terms of language and grammar; then, the students were asked to produce another writing piece of work on a similar topic to their first attempt. The study results indicated an increase in the total number of words, the unique words and the average word number per sentence in the improved version of their work (Can & Bardakci, 2022; Kallou & Kikilia, 2021). Overall, the results of the study were promising, indicating that ChatGPT could act as a language learning tool in L2 writing and contribute to a supportive learning environment, which is in line with the few relevant studies (Baskara & Mukarto, 2023; Yan, 2023). To the best of the researcher’s knowledge, the focus has not been put on multilingual classrooms, as students with a refugee/migrant background, who belong to socially vulnerable populations, have different language needs, face extra language difficulties and need extra help and support on the educators’ part to overcome language barriers and become efficient learners (Beacco et al., 2017; Dryden-Peterson, 2015; Kantzou et al., 2017).

Since the launch of ChatGPT in November 2022, there has been considerable debate about its impact on the field of education, as some researchers acknowledged its potential as a learning tool to support education, while others pointed out some drawbacks, highlighting that its use should be viewed with an amount of scepticism. Namely, ChatGPT can be abused and possibly hinder the learning process, as it can be used for cheating by directly providing answers or performing tasks instead of learners (Kim et al., 2023; Kohnke et al., 2023). Turnitin was recently upgraded to detect text written by Al. Simultaneously, it was mentioned that ChatGPT could not provide detailed and systematic feedback that could lead learners to improve their communication skills (Kim et al., 2023). Moreover, a critical stance is required on the learners’ part, as the information’s accuracy and the responses provided by ChatGPT should always be checked (Kim et al., 2023; Kohnke et al., 2023). To avoid possible abuse of the latest model of AI, it is strongly advised that a human educator is needed to supervise learners while employing ChatGPT in the educational context, as was the case of the present study. Krügel et al. (2023) suggested that relevant training should be provided to enhance users’ digital literacy skills.

Despite the ethical and societal impact of rapidly advancing technologies, such as ChatGPT, this study provides insights into its use as a foreign language learning tool to support writing in multilingual classrooms. It paves the way for more classroom-based research to exploit ChatGPT’s potential. However, there are some limitations of the present study that we need to consider, such as the relatively small sample of participants and the short duration of the teaching intervention, which does not allow us to generalize the findings of the research. After all, as mentioned above, this was a pilot study, which constituted the researchers’ first attempt to apply ChatGPT in multilingual classrooms and explore its feasibility. In this context, further
research is needed, including a more significant number of learners in different educational settings and language skills, to verify and prove the potential of ChatGPT in L2 language learning.

Overall, the promising findings of the current study testify that the integration of AI language models, such as ChatGPT, into the education setting has just begun, providing learners with dynamic and effective learning environments beyond traditional classrooms, which is consistent with relevant research (Ipek et al., 2023); they pointed out that performing high-level cognitive skills, such as information production, analysis, and synthesis through technological means has paved the way for a new paradigm, which education should take advantage of. However, to avoid the negative aspects of AI language models, they should be used in a responsible, ethical and wise manner, while users, educators and learners should hold a healthy amount of scepticism about their use (Aljanabi, 2023; Ipek et al., 2023). Finally, apart from the deficient and convenient sample, another limitation is that the study that uses self-reports runs the risk of socially desired responses and measurement bias (Lavidas et al., 2022).

References


McCarty, J. (2007). What is artificial intelligence?


