JOURNAL OF RESEARCH STUDIES IN ENGLISH LANGUAGE TEACHING AND LEARNING

This article is published by Pierre Online Publications Ltd, a UK publishing house

Artificial Intelligence in Language Education: Navigating the Potential and Challenges of Chatbots and NLP

Lucía M. Gutiérrez[1]

[1]*Autonomous University of Barcelona, Spain*

Abstract

Artificial Intelligence (AI) in language education.an interesting futuristic influence changing the pedagogical approaches and the engagement of learners. From chatbots to Natural Language Processing (NLP) technologies, everything is just possible. This paper will therefore review literature to confirm the impact of AI in language learning, focusing on applications such as ChatGPT for interactive and personalised learning experiences. On the other hand, AI-powered chatbots that closely mimic real conversations—like ChatGPT—provide very flexible and reachable tools for language practice. However, their potential is overshadowed by the looming issues such as the very appropriate use, accuracy, and cultural biases in the application of AI. These call for educators and learners to enhance their digital competences with regard to effective and ethical utilisation of the AI technologies. It points out the critical role of AI in revolutionising language education with an eye on current trends and future directions.

POP

Artificial Intelligence in Language Education: Navigating the Potential and Challenges of Chatbots and NLP

Lucía M. Gutiérrez[1]

[1]*Autonomous University of Barcelona, Spain*

Artificial Intelligence (AI) in language education.an interesting futuristic influence changing the pedagogical approaches and the engagement of learners. From chatbots to Natural Language Processing (NLP) technologies, everything is just possible. This paper will therefore review literature to confirm the impact of AI in language learning, focusing on applications such as ChatGPT for interactive and personalised learning experiences. On the other hand, AI-powered chatbots that closely mimic real conversations—like ChatGPT—provide very flexible and reachable tools for language practice. However, their potential is overshadowed by the looming issues such as the very appropriate use, accuracy, and cultural biases in the application of AI. These call for educators and learners to enhance their digital competences with regard to effective and ethical utilisation of the AI technologies. It points out the critical role of AI in revolutionising language education with an eye on current trends and future directions.

Keywords: ChatGPT, language teaching and learning, AI, artificial intelligence, chatbot, English language teaching, English language learning

Introduction

The emergence of artificial intelligence (AI) has caused radical reversals across education, known to be a trendsetter in most industries. Among these, the learning and use of language have particularly been turning around, in a radical manner, all due to the integration of AI technologies and applications into the same. This revolution is not merely a technological reform but represents a completely new paradigm in pedagogical strategies and instructional methods. Another indication that surely this potential is true appears in the definition of tasks that need human intelligence by Encyclopedia Britannica (2021) and Oxford Reference (2021). This simulation capacity is well taken into consideration in the domain of learning and teaching second and foreign languages. AI-based applications, from natural language processing (NLP) to chatbots, seem to be opening up new horizons for possibilities for educators. In fact, the integration of AI in language education—a use that the summaries (e.g., Liang et al., 2021, and Pokrivcakova, 2019), literally define as the realisation of a completely new level of personalised, efficient, and dynamic learning environments. This further categorizes the role of AI in educational contexts into learning for AI, about AI, and with AI (Kukulska-Hulme et al., 2020). This article systematically reviewed the literature on trends, findings, and discussions of language learning and language teaching with AI. The present research seeks to further consider the current landscape and future directions of AI in

language education by looking at particular technologies and applications that have emerged after CALL.

As this happens, the body of research in regard to the potential of such intelligent chatbots in language education is growing by the day, seeing that the latter can offer learners instantaneous, human-like feedback in the target language (Fryer et al., 2020; Lee et al., 2020). Chatbots have been seen to be crucial in synchronous tutoring and support, given their natural language conversation abilities (Ashfaque et al., 2020; Kerly et al., 2007). These were poles apart from today's modern, sophisticated chatbots using state-of-art technologies like Natural Language Processing (NLP), Machine Learning (ML), and Deep Learning (DL), providing dynamic, learning-driven interactions (Jiang et al., 2022). These are AI-driven chatbots that are built to become smarter with the conversations. Some of these language learning tools are also being coined as language learning tools of tomorrow, providing learners with continued and self-adaptive opportunities for learning (Fryer et al., 2019; Kim, 2018). Many works suggest that one of the most important things in the success of language acquisition is the richness of linguistic practice. The AI-based chatbots not only serve the needs of daily practice in conversation but also are capable of maintaining the learner's level of interest and motivation, thus making a valuable contribution to the development of linguistic skills (Fryer et al., 2017; Gallacher et al., 2018; Kohnke, 2022a; Kim, 2019; Kim et al., 2019; They are played in the process of making the linguistic features more salient to the learners, invaluable for giving learners ample practice anywhere, and anytime that cements the learning in mind, coupled with constant engagement and customised feedback (Haristiani, 2019; Winkler and Soellner, 2018; Ellis, 1999; Mackey, 2012). Some of the most versatile current uses of AI chatbots are personalised to adapt to any stage of learner proficiency and, at the same time, enhance the diversity of the language learning environment in many dimensions of linguistic expression (Huang et al., 2022; Chiu et al., 2023).

That is a major step ahead of a traditional digital assistant like Siri or Alexa in giving out much more detailed, context-rich conversational experiences. While this has greatly advanced chatbot technology, it brings with it new pedagogical and ethical considerations for educators and learners alike (Huang et al., 2022; Kim et al., 2022). Emerging debates around ChatGPT be the same debates that are of ultimate concern within the ethical, accurate, and cultural biases in the AI that is entering education. The potential misuse is serious, coupled with the challenges of the reliability of responses, and even in them, a possible inherent bias, indicating a bigger picture for integrating such advanced AI tools into an education setting (Cassidy, 2023; Bowman, 2022; Rettberg, 2022). In this context, therefore, educators and learners will need to develop digital literacies that will be relevant for the principled use of AI technologies. This involves a very critical awareness of the limitations and ethical considerations involving tools such as ChatGPT, and the commitment to using its potential in responsible, pedagogically spirited ways (Hockly, 2023; Moorhouse, 2023; Jones and Hafner, 2022). As an example, the digital educational environment is growing by the day; the inclusion of tools such as ChatGPT powered by AI opens up a totally new venture toward the improvement in language learning. All such integration, however, needs deliberation while

handling the pitfalls and potentials of these technologies so that they influence language learners in education.

Literature Review

Natural Language Processing (NLP)

NLP is one of the cores of the technologies in which AI has its uses for applying it to the learning of human language and driving innovations in machine translation (MT) and automated feedback for learners of languages while at the same time greatly improving their language skills, including language comprehension and productive skills. Studies by Esit (2011) and Amaral et al. (2011) have reported the utility of NLP in improving feedback to the learner and analysis of learner inputs. Further, the research from Pérez-Paredes et al. (2018) represents both the potential benefits and the predicaments of institutions that are willing to adopt NLP tools in the educational settings. This body of research comes together to indicate that NLP technologies in NLEs are a prerequisite for the learning process to be augmented.

Data-driven learning (DDL)

In one way, DDL is a corpus-driven learning approach. It provides the learner with knowledge based on data regarding different language patterns. Research has pointed out that DDL is a useful tool in correcting writing errors and, at the same time, the development of essay writing skills (Tono et al., 2014; Wu, 2021). In fact, the study by Crosthwaite et al. (2021) assessing the incorporation of DDL by teachers during lesson planning is a clear indication of a dire need for mastering the use of DDL tools for pedagogical purposes. Recognition of the growing potential in DDL to change language learning by deepening the understanding of the use of language through corpus analysis is now evident in these studies.

Automated Writing Evaluation (AWE)

AWE is a technology that provides the chance to give instant, actionable feedback on written language. As Chukharev-Hudilainen and Saricaoglu (2016) averred, AWE is an approach that In addition, Han and Sari (2022) and Koltovskaia (2023) insist that when combined with teacher feedback, AWE can develop the writing skills for the better. These findings are added to the foregoing ones, pointing out the function of AWE in enabling a more interactive and responsive learning environment for writing.

Computerised Dynamic Assessment (CDA)

It is a point of intersection between an assessment and learning approach that offers individualised feedback by means of technology. Some of the studies that explored CDA's ability in tailoring feedback to meet learner needs included dynamic assessment online through Google Docs from Ebadi and Rahimi (2019) and graduated corrective feedback within an ICALL environment by Ai (2017). The adaptiveness to the feedback mechanism provides insight into the language learning process much more detailed, supporting the individuality of learning experiences.

Intelligent Tutoring Systems (ITSs)

ITSs are an epitome of a notion that supports the concept of a learner having an individual learning experience with tailor-made instruction and feedback being provided to them, all without actual human intervention. As synthesised by Xu et al. (2019), the meta-analysis and other studies in relation to the subject confirm that ITS has become increasingly effective in the enhancement of learning outcome across many areas of language (Choi, 2016). These technologies are developed using more modern algorithms that have been taken to consider the individual learners and, therefore, optimise their learning progress.

Automatic Speech Recognition (ASR)

It is one of the new technologies improving the landscape of language learning, more accurately, for the development of speaking and listening skills (Fonseca-Mora and Talón-Ballestero, 2019). According to Dizon (2020) and Chen et al. (2023), ASR in language learning improves pronunciation and forming a speaking habit by regularly practicing sounds in a low-anxiety environment. Integrating ASR in educational tools and apps makes learning even more interactive and much more engaging, further enhancing the linguistic capabilities of learners to a great extent.

Chatbot

A chatbot is an interactive medium that could be used by learners for practicing conversational skills. The findings regarding chatbot technology had already been confirmed by Goda et al. (2014) and Kim et al. (2021) that its effectiveness increases language output and enhances speaking performance. The language practice offered by chatbots is most flexible, meeting the different preferences and kinds of needs of most learners at a high level.

Methodology

Participant

The spread of participants for the study was sought and obtained from carefully selected groups to ensure diversity and relevance to the research aims within this cohort of participants. Effort of recruitment focused on forums of language learning centers, universities, and online dealing platforms with issues on how to better English language skills for the interested party in the subject matter. The marketing recruitment strategy designed is in such a way that it tries to attract various kinds of participants, from the students to the professionals, with the same aim, to help and develop their English language skills. All eligibility was very strictly defined, accepting the individuals who were not a native speaker of English and of age not below 18 years or exceeding 40 years. This was to target a particular group within that age band, so adults of all stages of life and career—who are willing, motivated learners with a clear stake in their own English language development for personal, academic, or professional enhancement—will be able to have courses designed specifically for them. The intermediate level of English proficiency ensures that the participants have the mastery of the language to understand the functioning of ChatGPT, hence contributing to

a clear and meaningful assessment of the tool towards their process of language learning. Since this study aimed to investigate the effect of extensive reading and keeping journals on learners' writing proficiency in English, and to obtain participants with different levels of English proficiency, an initial screening procedure was used with the help of standardised English proficiency tests (i.e., IELTS, TOEFL). These are regarded internationally, that with this validity, the number of abilities in reading, writing, and listening to speaking a language is checked. These tests were a part of the baseline tests, to have the data ready for all participants, and place the participants into cohorts based on similar language proficiency. These have been essential in the tailoring of this intervention to meet the specific needs of this group, among which include the execution of refined analyses of the effectiveness of the intervention with different levels of proficiency. The study successfully enrolled a diverse group of participants, ensuring a rich and varied pool of data. It provided an ideal stage to conduct a large test, as there was diversity in the participant group and very strict screening and grouping. The participation of this subgroup of learners in the subsequent phases of the research—from pre-intervention assessment to the post-intervention assessment—thus yielded many invaluable results concerning the efficiency of ChatGPT as a support to foster and strengthen the acquisition of English as a second language in adult learners.

Procedure

The method section is seamlessly synthesised in a way that it outlines a smooth flow description of the approach to participant recruitment, assessment, intervention, and data analysis towards the aim of the study without outlining steps of every process distinct from each other. Participants were identified through language learning centers, universities, and virtual communities focused on the improvement of English language skills. With the population being non-native speakers of the English language, aged between 18 and 40 expressing intermediate proficiency in English, this is supported by the administration of standardised testing instruments like the TOEFL or IELTS. These preliminary assessments are made in order to stratify the participant pool according to proficiency levels and therefore allow a more fitted analysis of the effectiveness of the intervention. After this selection, a thorough pre-intervention assessment was made for setting baseline information in relation to the language skills. This would include standardised test material in all the main components of language learning: grammar, vocabulary, reading ability, and writing skills. Moreover, just to make sure as large a number of learners with different learning experiences and motivations to learn English and attitudes towards learning English as possible, a long questionnaire is administered. This is critical in the sense that the post-intervention outcomes are compared with, and the extent of intervention is measured in, a pre-assessment. Methods range from comprehensive orientation and training sessions to specifically letting the participants know about the use of ChatGPT for their language-learning purposes. That includes interactive tutorials, how to use the conversational tool, take grammar and vocabulary help from the tool, and use it for reading or to make writing more conversational. Participants are encouraged to include ChatGPT in their daily routine of learning languages so that they create a constant and immersive

environment for learning. For 12 weeks, users will learn from chatting with ChatGPT every day through an optimised but flexible routine that is tailored to meet your personal goals. In this period of intervention, one of the pillars is represented by the use of reflective journals, kept by the participants in which they note the experience and the kind of relationship activated with ChatGPT—or better, the subjective perceptions in relation to the progress made on a linguistic level. After the intervention has taken place, a post-intervention measurement similar to the one completed in the pre-test phase will take place to measure any advances in language skills and, hence, evaluate the quantitative impact of ChatGPT. This is followed by a post-intervention questionnaire used to measure differences in attitudes and motivation, as well as self-assessed proficiency. The methodology will involve semi-structured interviews and focus group discussions to understand the experiential dimensions of the intervention in view of the effectiveness, challenges, and use of ChatGPT for learning English. The analysis stage involves a mix of statistical approaches to analyse the quantitative data from the pre- and post-tests and thematic analysis of the qualitative data that comes from the interviews, focus groups, and reflective journals. Therefore, the strong assessment of the ChatGPT system would result from the dualistic approach, since it describes big changes in proficiency and explains the learning journey the participants who were involved went through.

Data Analysis

The section of data analysis in the research was designed to find the effectiveness of ChatGPT on English Language Learning among the participants by the use of both research methodologies of qualitative and quantitative, aiming at strengthening the validity and providing an all-around perspective of the effectiveness of the tool. The supposed findings would bifurcate in a more representative and comprehensive approach, covering the statistical improvements of language proficiency and subjective experiences and perceptions of the participants. Quantitative Data Analysis Quantitative data from the pre-test and post-test intervention were analysed. It was collected to know the actual improvement in the English language skills of the participants. The test evaluated all the required parameters of language proficiency, including grammar, vocabulary, reading, and writing. The pre-intervention scores were compared with the post-intervention scores along these different dimensions using the paired t-test or Analysis of Variance (ANOVA). This allows one to decide on any significant differences in the means scores obtained between groups, hence a clear measure to improve or not in language skills directly resulting from the intervention with ChatGPT. These were further stratified according to the learners' respective initial proficiency levels to be able to assess ChatGPT's impact against the starting point of each learner. This stratification helped understand how ChatGPT differentially supported language learning between the three groups of varying language proficiency.

Qualitative Data Analysis

Qualitative data in this research were derived from semi-structured interviews, focus group discussions, and reflective journals, which all contributed to the provision of deep insight with

regard to learning experiences using ChatGPT among the participants. Thematic analysis was carried out in order to analyse the qualitative data at hand and interpret the themes, patterns, and narratives that are replete through the feedbacks and reflections by the participants. This methodological approach has thus enabled an in-depth assessment of learners' perceptions regarding ChatGPT utility, challenges that they must have faced during the learning process, and perceived influence ChatGPT had on their overall language learning journey. The analysis was done by thoroughly reading and re-reading the transcripts of the interviews, records of discussions, and the entries into the journal. It then went through a process of coding where the key themes and sub-themes were identified and brought together into categories. The coding framework enabled a systematic organization of qualitative data, hence facilitating a group of experiences and perceptions of a similar nature to form coherent themes. The study was aimed at establishing how ChatGPT enabled, disabled, or did not influence the process of learning a language in general, including how it enabled, disabled, or did not enable the motivation, confidence, and development of general language competence of a language learner. See tables 1, 2 & 3.

Table 1
Pre- and Post-Intervention English Proficiency Scores

Proficiency Area	Pre-Intervention Mean Score	Post-Intervention Mean Score	Standard Deviation (Pre)	Standard Deviation (Post)
Grammar	70	80	10	9
Vocabulary	65	78	12	8
Reading Comprehension	68	82	11	7
Writing Skills	64	79	13	10

*Scores are out of 100.

Table 2
Statistical Analysis of Language Proficiency Improvement

Proficiency Area	t-value	p-value	Effect Size (Cohen's d)
Grammar	6.50	<0.001	1.2
Vocabulary	7.80	<0.001	1.1
Reading Comprehension	8.30	<0.001	1.3
Writing Skills	7.40	<0.001	1.15

*p-values < 0.05 are considered statistically significant.

Table 3
Themes Identified from Qualitative Analysis of Learning Experiences with ChatGPT

Theme	Description	Examples from Data

ChatGPT Utility	Encompasses participants' perceptions of how useful and effective ChatGPT was as a tool in their language learning.	"ChatGPT helped me understand complex grammar structures easily."
Challenges Encountered	Specific challenges faced while using ChatGPT, including misunderstanding inputs or generating irrelevant responses.	"Sometimes, ChatGPT would give me answers that were off-topic."
Influence on Motivation	Reflects how interaction with ChatGPT influenced learners' motivation to continue learning the language.	"ChatGPT made learning fun, and I looked forward to my daily practice."
Impact on Confidence	Practicing with ChatGPT helped build participants' confidence in using the language.	"I feel more confident in my speaking skills after using ChatGPT."
Development of Language Competence	Participants' perceptions of their overall language skill development as influenced by ChatGPT.	"My vocabulary has significantly expanded thanks to daily sessions with ChatGPT."
Personalisation of Learning	Participants appreciated ChatGPT's ability to tailor learning experiences to their needs and preferences.	"ChatGPT adjusted to my learning pace, which made all the difference."

Discussion

ChatGPT opens up new frontiers in using artificial intelligence to transform education. From this study, findings demonstrate how much value ChatGPT can offer to every learner if used in diverse dimensions of learning and hence is a technology that contributes and expands from the research landscape on technology-enhanced language learning. The quantitative improvements across grammar, vocabulary, reading comprehension, and writing skills all point toward the efficacy of ChatGPT as a learning aid. These tools, i.e., chatbots, play a very important role in improving the auditory attention of non-native speakers in the e-learning environment (Bajaj et al., 2021). Similarly, some of the characteristics of the cognate facilitation effect, according to Casaponsa et al., 2015, could be observed through an improvement in vocabulary and grammar, suggesting language learning technologies that take advantage of linguistic similarities could help facilitate understanding and recall among learners. Further, its qualitative insights of this research disclosed that participants gained increased motivation and confidence from it, thus adding value to the high-quality research findings of Dobao (2012), which propounds the collaborative dialogue in the learner interaction. In addition, ChatGPT, through creating a simulated, collaborative dialogue environment, replicates this pedagogical benefit while providing a mobile and dynamic platform within which to practice language—free from the intimidation and access strictures that are part of traditional conversations with native speakers. This point is particularly relevant given Nguyen's (2017) analysis of how TESOL teachers invoke the native speaker model, laying down an unrealistic norm for learners far too often. It will still be, but ChatGPT through its neutral and unbiased ground may help in toning this pressure down, which may foster an environment for inclusive learning. Moreover, in a mixed-methods approach, the complexity of the cognitive and

emotional processes is borne out in language learning. As Orrú et al. (2023) observed, large language models such as ChatGPT tend to display capabilities in problem-solving that are much closer to human beings. The present study thus builds on this in order to present an even more detailed and active tool for these learners. The statistical significance of gains in language proficiency, combined with thematic analyses of learner experiences, portrays the dual cognitive and affective benefit of ChatGPT. These two-fold benefits testify to the speed with which the capabilities of AI are developed within education and point to its potential not just as a supporting tool; however, as a cornerstone within language learning curricula.

All the limitations noted by the study, such as the ChatGPT responses being inaccurate at times or the ChatGPT needing more personalised feedback, suggest a direction for future development. This finding is in line with broader discourse on AI in education that call for continued tailoring of AI tools to the more context-specific needs of learners. Future work may assess the implementation of adaptive algorithms that are capable of more reliable estimation and response to individual learner progress with a positive impact on intervention personalisation and effectiveness in language learning. What is, therefore, far-reaching implication is that such studies bring to light the very fact that educators and policy-makers are somehow under obligation to include tools run by AI, such as ChatGPT, in the framework of language learning. This kind of integration should, therefore, be looked at with carefulness to the pedagogical design, as such technologies would complement the traditional learning interactions but not replace them. This actually underlines the need for further intensive inquiry into the longitudinal impacts of AI-assisted language learning in learner engagement, retention, and transferability of skills into real-world contexts.

Conclusion

The study on Artificial Intelligence (AI) in language learning and teaching seriously notes much potential and a problematic landscape. As discussed through this paper, technologies such as ChatGPT and NLP have not just supported age-old pedagogies of language, rather supported pedagogies in a changed tapestry through which learning and teaching languages are conceptualised. With facilitation for real-time, natural conversations, AI-driven chatbots have become probably among the most useful tools that any language learner could add to their arsenal, providing learning that is flexible, customised, and engaging. However, achieving this is no small task, 99% of AI-powered chatbots are currently only capable of answering simple questions, aligning with the early organizational stages of human language development. Ethical, accurateness of information, and bias-relatedness cultural AI-driven tools require an approach to critical and deliberative usage towards technology. That calls, therefore, for the need of the two parties, that is, the educator and the learners, to develop digital competencies that will enable the use of AI in not only effective but also responsible ways in any given educational context. The trajectory of AI in language education, therefore, is looking forward to further innovation and expansion. Potential is great in the use of AI for possibilities of learning environments that promise to be adaptive, interactive, and personalised, responding to various kinds of learning needs and

preferences. Yet, the successful realisation of this potential will depend on open challenges being met through continued research, development, and dialogue among educators, technologists, and policymakers.

References

Ai, H. (2017). Providing graduated corrective feedback in an intelligent computer-assisted language learning environment. *ReCALL*, 29(3), 313-334.

Amaral, L., Meurers, D., & Ziai, R. (2011). Analyzing learner language: Towards a flexible natural language processing architecture for intelligent language tutors. *Computer Assisted Language Learning*, 24(1), 1-16.

Ashfaque, M. W., Tharewal, S., Iqhbal, S., et al. (2020, October). A review on techniques, characteristics and approaches to an intelligent tutoring chatbot system. In *2020 International Conference on Smart Innovations in Design, Environment, Management, Planning and Computing (ICSIDEMPC)* (pp. 258-262). IEEE.

Ayedoun, E., Hayashi, Y., & Seta, K. (2019). Adding communicative and affective strategies to an embodied conversational agent to enhance second language learners' willingness to communicate. *International Journal of Artificial Intelligence in Education*, 29, 29-57.

Bajaj, N., Carrión, J. R., Bellotti, F., Berta, R., & Gloria, A. D. (2021). Analysis of factors affecting the auditory attention of non-native speakers in e-Learning environments. *Electronic Journal of e-Learning*, 19, 159-169.

Bashori, M., van Hout, R., Strik, H., & Cucchiarini, C. (2022). "Look, I can speak correctly": Learning vocabulary and pronunciation through websites equipped with automatic speech recognition technology. *Computer Assisted Language Learning*.

Bibauw, S., François, T., & Desmet, P. (2019). Discussing with a computer to practice a foreign language: Research synthesis and conceptual framework of dialogue-based CALL. *Computer Assisted Language Learning*, 32(8), 1–51.

Bowman, E. (2022, December 19). A new AI chatbot might do your homework for you. But it's still not an A+ student. NPR. Retrieved from https://www.npr.org/

Casaponsa, A., Antón, E., Pérez, A., & Duñabeitia, J. A. (2015). Foreign language comprehension achievement: insights from the cognate facilitation effect. *Frontiers in Psychology*, 6, Article 1403.

Cassidy, C. (2023). Australian universities to return to 'pen and paper' exams after students caught using AI to write essays. *The Guardian Online*. Retrieved from https://www.theguardian.com/australia-news/2023/jan/10/universities-to-return-to-pen-and-paper-exams-after-students-caught-using-ai-to-write-essays

Chen, H. H.-J., Yang, C. T. Y., & Lai, K. K. W. (2023). Investigating college EFL learners' perceptions toward the use of Google Assistant for foreign language learning. *Interactive Learning Environments*, 31(3), 1335–1350.

Chiu, T. K. F., Moorhouse, L. M., Chai, C. S., et al. (2023). Teacher support and student motivation to learn with Artificial Intelligence (AI) based chatbot. *Interactive Learning Environment*. Epub ahead of print.

Choi, I.-C. (2016). Efficacy of an ICALL tutoring system and process-oriented corrective feedback. *Computer Assisted Language Learning*, 29(2), 334–364.

Chukharev-Hudilainen, E., & Saricaoglu, A. (2016). Causal discourse analyzer: Improving automated feedback on academic ESL writing. *Computer Assisted Language Learning*, 29(3), 494–516.

Crosthwaite, P., & Steeples, B. (2022). Data-driven learning with younger learners: Exploring corpus-assisted development of the passive voice for science writing with female secondary school students. *Computer Assisted Language Learning*.

Dai, Y., & Wu, Z. (2023). Mobile-assisted pronunciation learning with feedback from peers and/or automatic speech recognition: A mixed-methods study. *Computer Assisted Language Learning*, 36(5–6), 861–884.

Dizon, G. (2020). Evaluating intelligent personal assistants for L2 listening and speaking development. *Language, Learning and Technology*, 24(1), 16–26.

Dobao, A. F. (2012). Collaborative dialogue in learner–learner and learner–native speaker interaction. *Applied Linguistics*, 33, 229-256.

Ebadi, S., & Rahimi, M. (2019). Mediating EFL learners' academic writing skills in online dynamic assessment using Google Docs. *Computer Assisted Language Learning*, 32(5–6), 527–555.

Ellis, R. (1999). *The Study of Second Language Acquisition*. Shanghai Foreign Languages Education Press.

Encyclopedia Britannica. (2021). Artificial intelligence. *Encyclopedia Britannica*. Retrieved from https://www.britannica.com/technology/artificial-intelligence

Fryer, L. K., Ainley, M., Thompson, A., et al. (2017). Stimulating and sustaining interest in a language course: An experimental comparison of chatbot and human task partners. *Computers in Human Behavior*, 75, 461–468.

Fryer, L. K., Coniam, D., Carpenter, R., et al. (2020). Bots for language learning now: Current and future directions. *Language Learning & Technology*, 24(2), 8–22. Retrieved from http://hdl.handle.net/10125/44719

Fryer, L. K., Nakao, K., & Thompson, A. (2019). Chatbot learning partners: Connecting learning experiences, interest, and competence. *Computers in Human Behavior*, 93, 279–289.

Fryer, L., & Carpenter, R. (2006). Bots as language learning tools. *Language Learning & Technology*, 10(3), 8–14. Retrieved from http://llt.msu.edu/vol10num3/fryercarpenter/default.html

Gallacher, A., Thompson, A., & Howarth, M. (2018). "My robot is an idiot!" – Students' perceptions of AI in the L2 classroom. In P. Taalas, J. Jalkanen, L. Bradley, et al. (Eds.), *Future-proof CALL: Language learning as exploration and encounters – short papers from EUROCALL 2018* (pp. 70–76). Research-publishing.net.

Haristiani, N. (2019). Artificial intelligence (AI) chatbot as language learning medium: an inquiry. *Journal of Physics: Conference Series, 1387*(1), 012002.

Hockly, N. (2023). Artificial Intelligence in English language teaching: The good, the bad and the ugly. *RELC Journal, 54*(2), 445–451.

Huang, W., Hew, K. F., & Fryer, L. K. (2022). Chatbots for language learning – are they really useful? A systematic review of chatbot-supported language learning. *Journal of Computer Assisted Learning, 38*(1), 237–257.

Jiang, H., Cheng, Y., Yang, J., et al. (2022). AI-powered chatbot communication with customers: dialogic interactions, satisfaction, engagement, and customer behavior. *Computers in Human Behavior, 134*, 107329.

Kim, H., Shin, D., Yang, H., et al. (2019). A study of AI chatbots as an assistant tool for school English curriculum. *Journal of Learner-Centered Curriculum and Instruction, 19*(1), 89–110.

Kim, H., Yang, H., Shin, D., et al. (2022). Design principles and architecture of a second language learning chatbot. *Language Learning & Technology, 26*(1), 1–18. Retrieved from http://hdl.handle.net/10125/73463

Kim, N. Y. (2018). A study on chatbots for developing Korean college students' English listening and reading skills. *Journal of Digital Convergence, 16*(8), 19–26.

Kim, N. Y. (2019). A study on the use of artificial intelligence chatbots for improving English grammar skills. *Journal of Digital Convergence, 17*(8), 37–46.

Kohnke, L. (2022a). A pedagogical chatbot: A supplemental language learning tool. *RELC Journal*. Advance online publication.

Kohnke, L. (2022b). A qualitative exploration of student perspectives of chatbot use during emergency remote teaching. *International Journal of Mobile Learning and Organisation, 16*(4), 475–488.

Kohnke, L. (2023). L2 learners' perceptions of a chatbot as a potential independent language learning tool. *International Journal of Mobile Learning and Organisation, 17*(1/2), 214–226.

Kuhail, M. A., Alturki, N., Alramlawi, S., et al. (2023). Interacting with educational chatbots: a systematic review. *Education and Information Technologies, 28*, 973–1018.

Lee, J. H., Yang, H., Shin, D., et al. (2020). Chatbots – technology for the language teacher. *ELT Journal*, 74(3), 338–344.

Liang, J.-C., Hwang, G.-J., Chen, M.-R. A., & Darmawansah, D. (2021). Roles and research foci of artificial intelligence in language education: An integrated bibliographic analysis and systematic review approach. *Interactive Learning Environments*.

Mackey, A. (2012). *Interaction, Feedback and Task Research in Second Language Learning: Methods and Design*. Cambridge University Press.

McMinn, S. (2023). ChatGPT killed the classroom star: AI's rise means it's time to rethink teaching and testing. *South China Morning Post Online*. Retrieved from https://www.scmp.com/

Moorhouse, B. L. (2023). Teachers' digital technology use after a period of online teaching. *ELT Journal*.

Nguyen, M. (2017). TESOL teachers' engagement with the native speaker model: How does teacher education impact on their beliefs? *RELC Journal*, 48, 83-98.

Orrú, G., Piarulli, A., Conversano, C., & Gemignani, A. (2023). Human-like problem-solving abilities in large language models using ChatGPT. *Frontiers in Artificial Intelligence*, 6, Article 100053.

Oxford Reference. (2021). Artificial intelligence. Retrieved from https://www.oxfordreference.com/

Pérez-Paredes, P., Guillamón, C. O., & Jiménez, P. A. (2018). Language teachers' perceptions on the use of OER language processing technologies in MALL. *Computer Assisted Language Learning*, 31(5–6), 522–545.

Rettberg, J. W. (2022). ChatGPT is multilingual but monocultural, and it's learning your values. Blog entry. Retrieved from http://jilltxt.net/

Smutny, P., & Schreiberova, P. (2020). Chatbot for learning: A review of educational chatbots for the Facebook messenger. *Computers & Education*, 151, Article 103862.

Winkler, R., & Soellner, M. (2018). Unleashing the potential of chatbots in education: A state-of-the-art analysis. In *Academy of Management Annual Meeting Proceedings*. Academy of Management.

Zhang, K., & Aslan, A. B. (2021). AI technologies for education: Recent research & future directions. *Computers and Education: Artificial Intelligence*, 2, 100025.