LLM-Based Chatbots in Language Learning

Panagiotis Panagiotidis
Aristotle University of Thessaloniki, Greece

Abstract

Efforts to utilize AI in education, and especially in language education, have their roots in the 60s with the appearance of the first rule-based systems. However, recent advances in Artificial Intelligence (AI) and more specifically the introduction of ChatGPT, have given a new perspective to language learning. The integration of AI, natural language processing, and Machine Learning has enabled adaptive learning environments tailored to individual learners' needs and led to a new generation of advanced tutoring systems and chatbots, able to offer personalized and customizable learning experiences to enhance language learning by increasing learner autonomy, engagement, motivation, and effectiveness. At the same time, research on Large Language Models revolutionized chatbot capabilities, making them integral tools in language learning. Commercial language learning online platforms experienced similar advancements, incorporating AI tools that offer enhanced possibilities for personalized language learning (PLL). This paper examines how the introduction of Large Language Models (LLM) and conversational AI has revolutionized the capabilities of chatbots, introducing the transformative potential of AI technologies to language learning and enabling them to enhance learning experiences, to increase engagement and improve their language proficiency. It also presents the existing specialized online language learning platforms and analyzes the areas in which artificial intelligence tools are currently used in language teaching and the benefits they have brought to users. The paper also discusses the problems, challenges and implications related to LLMs, such as ethical issues, potential biases and privacy concerns that need to be addressed, as well as the areas in which future research must focus, such as the pedagocical exploitation of AI tools and the most effective ways, strategies and pedagogical approaches and methodologies to blend AI-driven learning with traditional instruction, so that the use of artificial intelligence in language education becomes beneficial.

Keywords: artificial intelligence, iCALL, personalized language learning, LLM, large language models, natural language processing, machine learning, chatbots, educational technology. adaptive learning, language learning

1. Introduction: From CALL to iCALL

The need for individualized learning, i.e. Personalized Language Learning (PLL), has always been a target for language education experts. Early CALL (Computer Assisted Language Learning) systems (1960s-1980s) as well as Communicative CALL applications (1980s-1990s) certainly did not offer such characteristics. In the last generation of Integrative CALL (1990s-2000s) the integration of multimedia and internet resources allowed real-time communication with native speakers, fostering greater immersion and practical use of language skills and granted users to have a certain degree of autonomy, thus approaching the notion of Personalized Language Learning.

However, the concept of PLL really began to be approached more meaningfully with the advent of Artificial Intelligence (AI) and adaptive learning. Adaptive learning is an educational approach that combines CALL with artificial intelligence to provide personalized resources and activities tailored to the specific needs of each student (Wang et al., 2023). According to Zerkowska (2024), adaptive learning is a technological method for implementing personalized learning, which focuses on continuously adapting the material to match the learner's skills.

Advances in Artificial Intelligence, Natural Language Processing (NLP), and Machine Learning (ML) at the beginning of the 21st century gave a boost to research on adaptive learning applications and led to the emergence of iCALL (Intelligent Computer-Assisted Language Learning), the most recent evolution -or, for some, a subfield- of CALL (Kannan & Munday, 2018). iCALL systems use artificial intelligence to provide more personalized and adaptive learning experiences. Unlike the "one-size-fits-all" approach of the traditional CALL systems, iCALL adapts learning activities to the individual needs of learners. iCALL applications are also capable of providing personalized instructions and feedback, of correcting grammatical errors or translating texts as well as creating language content (Choi, 2016) and of creating interactive and immersive AR/VR learning environments (De la Vall & Araya, 2023) thus managing to provide a more authentic and engaging learning experience for language learners by engaging them in real-world contexts and tasks. Some examples of recent iCALL applications are: Listening Hacked (Vu et al, 2022), CIPW ICALL System (Chen et al, 2022), AutoTutor (Fang et al, 2022), French TPRAI motion sensing teaching system (Huang & Wang, 2021), ViewTones / Trip2China (Wang, 2024).

In an attempt to promote the development of iCALL applications for language learning, Hassani et al., (2016) proposed the iCALL framework, which outlines rules and principles for creating and using AI language learning systems. According to Bibauw et al., (2019), the iCALL framework offers guidelines for developing AI language learning tools that enhance the effectiveness of the learning procedure as well as user satisfaction and engagement.

Today, ongoing research on iCALL is focusing on a variety of topics, such as PPL, Tutoring systems, AR/VR environments, Automated Writing Evaluation and, of course, chatbots, with the common goal to create personalized and adaptable learning experiences and to promote learner engagement, motivation, and effectiveness thus enhancing language learning (Woo & Choi, 2021).

2. Evolution of Chatbots

Chatbots -or conversational agents or simply bots-, are software or AI applications that use natural language to interact with users (Bibauw et al., 2019; Mavropoulou & Arvanitis, 2023). Chatbots can perform conversations in any specific topic via text or audio and respond to users with intelligent responses in natural language (Haristiani, 2019; Kim et al., 2022; Son et al., 2023). In their more recent form based on conversational AI, chatbots can function as intelligent tutors: they can chat with users, find relevant educational material, or provide them feedback on questions they ask etc. (Belda-Medina & Calvo-Ferrer, 2022; Mageira et al., 2022).

Chatbot technology has rapidly evolved over the past decades, primarily due to recent advancements in NLP and Machine Learning (Caldarini et al., 2022; Luo et al., 2022). The first attempts to create Chatbots started in 1966 with the presentation of ELIZA by MIT (Weizenbaum, 1966) and later the -improved in structure and emotional reactions- PARRY in 1972. These two chatbots operated automatically, without the ability to understand the meaning of the input. ELIZA utilized pattern matching and substitution techniques to mimic conversation, relying on scripted responses without comprehending the context or meaning behind the words. PARRY employed similar principles but incorporated more complex rules to simulate paranoid behavior (Colby et al., 1971).

Initially, chatbots were text-based and depended on preset Q&A scripts, leading to predictable responses and interactions that did not seem natural by human standards (Belda-Medina & Calvo-Ferrer, 2022). In the following years, thanks to the Introduction of NLP, chatbots with improved performance were introduced: Racter in 1984 (ELMCIP, 2023), Jabberwacky in 1988 (Singh & Thakur, 2020) and Dr. Sbaitso, the first application of text-to-speech technology, in 1990 (Deryugina, 2010). These chatbots were rule-based systems that evolved to incorporate basic NLP and were capable of processing and generating human language to some degree. By using advanced knowledge bases and context-aware algorithms that enable relevant responses, they succeeded in making the simulation of a casual human conversation more convincing. However, these systems still lacked the capability to deeply understand context.

In the late 90s, the incorporation of statistical methods and early Machine Learning models enhanced chatbots' capacity to manage more diverse conversations by learning from data rather than depending solely on predefined rules. An important step in chatbot evolution was ALICE, a sophisticated rule-based system introduced in

1995, which used its own Artificial Intelligence Markup Language (AIML) as well as pattern matching and predefined rules to simulate conversation (Wallace, 2009).

In the early 2000s, significant advancements in NLP lead to the presentation of chatbots that can understand and generate human language, thus offering a more natural interaction with users. In 2001, ActiveBuddy was introduced and in 2008, Cleverbot, which in 2011 became the first chatbot to successfully pass the Turing test, earning a 59.3% human-like rating (Wikipedia, 2024). In this period, the appearance of intelligent voice assistants also begins. In 2010 Apple introduced Siri which could perform tasks via voice commands, marking the integration of AI into consumer devices. In 2011 IBM introduced Watson (High, 2012), which used the power of AI in understanding and processing complex questions. In 2012, Google introduced Google Now, which could convert voice inputs into search results, and in 2014-2016 Microsoft introduced XiaoICE, known for its emotional intelligence and ability to maintain long-term interactions with users, and Tay for interaction with users in Twitter (Dam et al., 2024).

Towards the same direction, several chatbots have also been developed for language learning. The earlier mentioned Cleverbot, which could adapt its replies based on users' type (Haristiani, 2019), CLIVE, an intelligent chatbot for conversational language practice (Zakos & Capper, 2008), Gengobot, a grammar dictionary application based on a chatbot for Japanese language learning (Haristiani et al., 2019), BookBuddy, a system designed to help children choose books suitable for their English language proficiency level (Ruan et al., 2019) and CSIEC, which was capable of generating communicative responses based on user input (Jia, 2009), are among the best known and widely used.

To summarize, from their first generation to today, chatbots have traveled an evolutionary path that transformed them from simple programs without intelligence that return answers based on preset Q&A scripts, to real interlocutors with emotional intelligence able to maintain long-term interactions with users, and to provide them relevant feedback on questions they ask.

3. Conversational AI

After 2010, the advent of deep learning and advanced NLP models revolutionized chatbots. AI and ML integration enhanced chatbot capabilities, enabling more natural and context-aware interactions. Conversational AI allowed chatbots to comprehend context, intent and sentiment, resulting in more natural interactions.

Conversational AI refers to a set of technologies -AI, NLP, ML, Speech Recognition and others- which gives computers the ability to understand, process, and generate human language and thus engage in human-like conversations. This technology is used by chatbots, virtual assistants, and other interactive platforms that mimic human conversation. Conversational agents typically operate online and can interact with users through text, voice, graphics, virtual gestures, or haptic feedback (Belda-

Medina & Calvo-Ferrer, 2022), engaging seamlessly in natural language exchanges with both humans and other agents.

The first generation of this technology was able to perform short dialogues for specific tasks, such as playing music (Alexa) or retrieving information (Cortana) (Mageira et al., 2022). Early Intelligent Personal Assistants (IPAs) were able to provide standardized answers to common questions, but later, thanks to advances in Natural Language Understanding (NLU), neural networks and deep learning techniques, new sophisticated programs were developed. The introduction of Amazon's ALEXA (2014), Microsoft's Cortana (2016), Google Assistant (2016) and Samsung's Bixby (2017) made possible the use of Conversational Agents in everyday tasks on users' computers or mobile devices (de Barcelos Silva et al., 2020). The maturation of this technology led to its widespread adoption as it proved its effectiveness in common tasks such as search and information retrieval. Conversational AI has been integrated into language learning platforms in various ways, including the provision of personalized learning content, tutoring and assistance, translation, automatic writing assistance, language practice and conversation. Some examples of Conversational AI in educational contexts are: Khan Academy's Khanmigo (Tutor and teaching assistant), Carnegie Learning's MATHia and Pearson's Aida Calculus (Mathematics), Quizlet's Learn Mode (personalized study plan creator), Kahoot (AI-powered Question Generator feature), Ginger and ProWritingAid (feedback provision), IBM Watson Tutor (ITS), Microsoft Copilot for Microsoft 365, and of course a variety of Language Learning Applications (Grammarly, Notion AI, Asasara Bot, LanguaTalk, HelloTalk, ChatClas, Elsa Speak and several online language learning platforms, see ch. 7).

To summarize, Conversational AI led to the transformation of chatbots from basic, "rule-based" systems to sophisticated, intelligent agents capable of producing natural text and engaging in coherent context-aware interactions. By integrating advanced NLP, Machine and Deep Learning techniques, Conversational AI has significantly expanded the capabilities and applications of chatbots across various sectors, creating huge expectations for what they can achieve in the future.

4. The Rise of Large Language Models (LLMs)

At the same time, ongoing research on Linguistic Models has led to the emergence of Large Linguistic Models (LLMs) thus transforming the field of natural language processing and drastically enhancing the development and capabilities of chatbots. Dam et al. (2024) essentially distinguish two generations of chatbots: The 1st generation which concerns chatbots that were developed before the appearance of LLMs (Pre LLM) and those that utilized LLMs when they became available.

Language Models (LMs) are Machine Learning models trained to perform probability distributions on words, which enables them to understand and generate human language (Devlin et al., 2018). Large Language Models (LLMs) represent significant

progress in NLP as they are advanced language models with billions to trillions of parameters and exceptional learning capabilities (Chen et al., 2021; Zhao et al., 2023).

LLMs became widely known to the public when OpenAI introduced GPT (Generative Pre-trained Transformer) in 2018. Very soon, larger, more sophisticated language models were presented, such as GPT-2 (Radford et al., 2019), GPT-3 (Brown et al., 2020), GPT-3.5 (Ye et al., 2023), and GPT-4 (OpenAI, 2024) all language prediction models based on neural networks. At the same time, similar LLMs were also presented by other companies, research groups or experts in education: BERT (Devlin et al., 2018), PaLM (Chowdhery et al., 2022), LLaMA (Touvron et al., 2023), XLNet (Yang et al., 2019), T5 (Raffel et al., 2020), RoBERTa (Liu et al., 2019), and BLOOM, an open-source project offering a transparently trained multilingual language model specifically designed to support 46 natural languages and 13 programming languages (Scao et al., 2022).

The development of LLMs is based in Deep Learning techniques, particularly the "Transformer" architecture (the Vanilla Transformer model) and the underlying attention mechanism proposed by Vaswani et al., (2017), who can capture contextual relationships far more effectively than traditional neural networks. Transformers harness self-attention mechanisms, enabling the model to evaluate the importance of individual words within a sentence.

LLMs are -in the first stage- trained on massive amounts of text data, in order to learn linguistic patterns, syntax and semantics. As a result, LLMs can generate human-like text, have conversations, answer questions and efficiently complete other language-related tasks (Kasneci et al., 2023). This stage is followed by a fine-tuning stage on specific tasks or domains, such as text generation, translation, and summarization, which further improves performance on a wide range of language tasks (Wei et al., 2021).

Another important feature of LLMs, in-context learning, firstly introduced in GPT-3, allows the model to understand the prompts given by the users and to generate realistic responses without any additional training (Brown et al., 2020).Thus, LLMs can generate contextually relevant text in various forms, such as articles, reports, posts or even poems, enhancing content creation and improving translation or helping in cross-cultural communication. In users' questions (prompts), LLMs provide accurate, contextually appropriate responses, in various sectors such customer service, healthcare, journalism and education. Moreover, LLMs can perform summarization tasks, produce programming code, or help in information retrieval, reducing cognitive overload for researchers and professionals and preparation time for educators. When properly trained on multilingual data, LLMs can handle and produce text in multiple languages, although their proficiency varies depending on the quality and quantity of data available for each language.

However, LLMs actually lack true semantic understanding, as all the answers they give are based on the data they have been fed with during their training. This is why they occasionally generate hallucinations, resulting in false, incorrect or illogical answers (Bang et al., 2023; Wu et al., 2023; OpenAI, 2024) which shows that attention, critical thinking and control are required from users.

LLMs also face challenges: supercomputers needed to run artificial intelligence programs, require huge amounts of water and electricity (Sisson, 2024). Privacy and ethical issues also arise, as LLMs may inadvertently violate individuals' privacy by reproducing sensitive information or propagate false/fake information found in large data sets, leading to biased results. Furthermore, as understanding how these models arrive at specific results is often opaque, the interpretability of LLMs remains a complex challenge (Kasneci et al., 2023).

Large Language Models signify exponential progress in natural language processing, offering innovative capabilities in text generation, translation, understanding, and the development of sophisticated chatbots. Thus, they can have significant practical implications for education. In this paper, only their contribution to the field of language learning is examined, which is further analyzed in ch.6. As LLMs are accessible to users through chatbots, it is useful to first introduce the most well-known LLM Chatbots.

5. LLM-based Chatbots

ChatGPT: ChatGPT, developed by OpenAI, was first released in November 2022 (ChatGPT-3.5), based on the capabilities of its predecessors in the GPT series. ChatGPT-3.5, which could generate human responses in conversational environments, was trained on massive text datasets using Deep Learning techniques, such as transformer models and reinforcement learning (Chowdhery et al., 2023). The release of ChatGPT was a major advance in AI, as for the first time LLMs became useful in everyday interactions for practical applications. ChatGPT's ability to provide relevant and realistic responses on a variety of topics immediately ensured its widespread recognition. Its recent improvement to the GPT-4 version in March 2023, made it more reliable and flexible, further enhancing the model's capabilities (reportedly handling 1.76 trillion parameters). From its evolution, ChatGPT has expanded its applications beyond simple chat to more complex tasks such as composing emails, creating content, and providing detailed information on a wide range of topics. Today, ChatGPT can engage in free-form dialogue and provide detailed responses while avoiding inappropriate or harmful content (Obaidoon & Wei, 2024) and stands as a robust conversational agent, widely used in various industries, reflecting rapid developments in AI technology and increasing integration into everyday life. As of September 2023, ChatGPT also gained the ability to search the web for up-to-date content (Al Jazeera, 2023), further enhanced by an updated knowledge base containing information up to April 2023 (Johns, 2023).

Google Bard / Gemini: Bard was introduced in February 2023 as Google's response to the impact that the introduction of ChatGPT had on search engines (Zhang et al., 2023). BARD is a publicly accessible user interface (chatbot) for an LLM designed for collaborative generative AI, aiming to combine external knowledge from the Internet with the internal capabilities of language models to deliver high-quality responses (Ortiz, 2024a). It utilizes Google's LaMDA (Language Models for Dialogue Applications) conversational AI model, and new techniques such as grounded learning to improve its information quality (Pichai, 2023). At the time of its release, BARD was theoretically superior to ChatGPT, as it could search for up-to-date information while ChatGPT had access only to information published before 2021 (Hetler, 2023). Earlier this year, Google rebranded BARD to 'Gemini' (Crider, 2024).

Bing Chat / Copilot. Soon after Google announced BARD, Microsoft launched Bing Chat (Condon, 2023), at the beginning in limited preview testing, to gather user feedback for improvement (Mehdi, 2023). It is a search engine feature driven by GPT4 that enables users to interact with an AI chatbot during their searches in Bing Search Engine. Bing Chat is powered by Prometheus, an AI model developed by Microsoft with over 200 billion parameters and trained on public conversation data and fine-tuned using human feedback (Obaidon & Wei, 2024) able to provide conversational search experiences. Thanks to its integration with Microsoft's search engine, Bing Chat was already able to provide live internet access and citation-supported responses upon its release (Freeman-Mills, 2023; Conway, 2023), as well as an advanced interface to provide a personalized interaction based on the user's query (Simplilearn, 2023). Later that year, Microsoft, trying to respond to the huge appeal of ChatGPT which had already passed 100 million users, decided to rebrand Bing Chat, renaming it Copilot (Warren, 2023). However, Copilot -particularly GitHub Copilot- is primarily considered an AI-powered code assistant rather than a traditional LLM chatbot. It uses LLM technology to understand and generate code snippets and provide programming assistance, but its primary function is to assist developers in writing code rather than engaging in general conversational interactions (Ortiz, 2024b).

Claude: Claude 1.0 was launched by Anthropic research lab in March 2023 (Anthropic, 2023), alongside the streamlined -a bit faster and lighter- version Claude Instant 1.1, aiming at providing helpful and honest conversational AI experiences using advanced LLM techniques (Dam et al., 2024). This is accomplished using constitutional AI, a fine-tuning process developed by Anthropic researchers (Lozić & Štular, 2023). More advanced versions, such as Claude 2 and Claude Instant 1.2, Claude 2.1 and Claude 3 Opus, followed (Gillham, 2023) until June 2024, when Anthropic launched Claude 3.5 Sonnet, which showed significantly enhanced performance on benchmarks compared to the larger Claude 3 Opus. A notable feature of Claude (2) was its extensive context window of 100,000 tokens, or approximately 75,000 words. This capacity allowed users to conduct thorough analyses of extensive documents (Bai et al., 2022). By default, Claude does not have the ability to browse

the internet. However, with ChatLabs AI, users can enable internet access for Claude 3 AI to perform web-related tasks (Chatlabs, 2023).

Ernie Bot: Ernie (Enhanced representation through knowledge integration) Bot is an advanced conversational agent launched by Baidu on March 2023 (Motlagh et al., 2023). It is designed to understand and generate human-like text, facilitating interactions through natural language processing. Ernie Bot, like other AI chatbots, can perform tasks such as answering questions, providing information and engaging in dialogue with users. To this end, Ernie has been trained on a vast amount of data, including search and image data, trillions of web pages, billions of voice data and a knowledge graph of 550 billion facts (Yang, 2023). Ernie's development aligns with Baidu's broader AI initiatives, integrating sophisticated Machine Learning techniques to enhance its conversational abilities and overall performance in various applications. Baidu plans to integrate Ernie into its search engine (Yang, 2023), thanks to its ability to read text in various Chinese dialects (Rudolph et al., 2023), as well as into other products, such as autonomous vehicles (Huang, 2023).

Apart from these major LLM Chatbots, many other chatbots, such as Deep-Mind's Sparrow (Rodriguez, 2023), xAI's Grok (Milmo, 2023), BlenderBot (Zhang, 2023), are also under development. It is obvious that there is ongoing research and rapid progress based on interdisciplinary collaboration in an effort to develop more effective and trustworthy chatbots and to balance innovation with ethical issues that are crucial to ensuring the reliability of chatbots.

LLM chatbots and especially ChatGPT, are now widely used in language learning. Several research proved their effectiveness and the usefulness in EFL (Mubaroq et al, 2024; Songsiengchai et al, 2023; Xiao & Zhi, 2023; Farhi et al, 2023), especially in the improvement of writing skills, reading, vocabulary and other language learning tasks. Obaidoon & Wei (2024) made a thorough comparison of ChatGPT, Claude and Bard, in evaluating Chinese writing which also proved the value they can add to the language learning process.

Comparing LLM chatbots for language learning applications is not an easy task as each one seems to have its own strengths. Comparative studies (Chang et al. 2024; Dam et al. 2024; Mistelbacher, 2024; Obaidoon & Wei, 2024), found small discrepancies regarding the effectiveness of these chatbots in everyday tasks and concluded that no model can perform perfectly in all kinds of work. Most of the available literature on the use of LLM chatbots in language education concerns ChatGPT, as it was the first to be made publicly available. It is a common conclusion that ChatGPT excels in natural language understanding and generation. It supports native voice interaction, allowing users to engage in spoken conversations, and offers enhanced analytical functionalities, helping with grammar and vocabulary exercises. On the other hand, Google's Gemini is designed for real-time conversational scenarios, making it ideal for interactive language practice. Gemini accepts various types of input (text, images, audio, and video) providing an immersive learning experience allowing learners to

engage with diverse content formats, enhancing comprehension and retention. Copilot can be beneficial for learners focusing on writing and editing skills, as it is integrated in Microsoft Office and Bing applications. Finally, Claude's extensive context handling may give it an advantage for in-depth textual analysis.

6. LLM Chatbots & language learning

LLM Chatbots offer a variety of possibilities that can enhance -or even revolutionize- the language learning process. Several studies have explored the potential implications of chatbots in language learning (Kim, 2020; Chuah & Kabilan, 2021; Jeon, 2021; Huang et al., 2022; Atlas, 2023). Son et al. (2023) distinguished seven categories -which are interconnected and may be combined based on the particular use per case-, into which AI technologies and applications for language education could be classified: NLP, automated writing evaluation (AWE), automatic speech recognition (ASR), computerized data-driven learning (DDL), intelligent tutoring systems (ITSs), dynamic assessment (CDA), and chatbots. In their research, focusing on how AI was integrated into language education, Huang et al. (2022) recorded a variety of AI applications able to help students in improving all four language skills and highlighted the importance of Automatic Writing Evaluation (AWE), Intelligent Tutoring Systems (ITS) and Personalized systems for language learning. They concluded that AI applications present a great potential in providing personalized learning experiences, in enabling immediate adjustment and offering opportunities to practice the target language, allowing students to learn anywhere and anytime. Obaidoon & Wei (2024) research findings also indicate a promising future for AI in education, with LLM chatbots nearing human level proficiency in their evaluations in terms of scoring. Since the exploitation of AI in language education is still a relatively new challenge, the available literature is difficult to be categorized. Summarising from the most cited research of the past four years, the main areas where LLM chatbots can be useful in language learning are briefly presented below.

LLM Chatbots can act as conversation partners offering real-time interaction thus allowing students to practice speaking and writing in a foreign language, understand idiomatic expressions and cultural references and improve their fluency and confidence in using the language (Brown et al., 2020).

The ability to handle multiple languages is also very important. This makes LLM chatbots capable of providing translation assistance in real time, thus helping students with reading comprehension exercises and clarifying the meaning of unknown words and phrases (Lewis et al., 2020). They also allow students to work with parallel texts in different languages. This can help them recognize differences in syntax, vocabulary and idiomatic usage, enhancing their understanding of target language structure and language fluency (Conneau et al., 2020).

A major asset of LLM Chatbots is their ability to serve language students by providing explanations and lessons on grammar and vocabulary, summaries, paraphrases and

explanations of complex texts, helping students with comprehension (Raffel et al., 2020) or by exposing them to authentic and natural language input through text and audio (Kim, 2020). They can also create customized language exercises for all four language skills, appropriate for the learner's level of knowledge. In addition, they can design personalized learning paths that adapt to the individual's needs (Xu et al., 2020) and create personalized learning materials (Pokrivcakova, 2019). Finally, LLM chatbots can be useful in different pedagogical approaches, such as Content and Language Integrated Learning (CLIL), i.e. the simultaneous learning of content (e.g. a scientific topic) and foreign languages (Mageira et al., 2022).

LLM chatbots are also very capable of providing writing assistance. They can significantly enhance academic writing and improve both the quality of content and individual writing skills by correcting punctuation, spelling, syntax and grammatical errors, promoting, at the same time, students' self-regulation and autonomy (Pokrivcakova, 2019). They can propose variations of a text, such as word choice or sentence structure or even stylistic variations, enhancing the quality and fluency of the learner's writing. They can also compose different types of text such as essays, stories, or poems (Kasneci et al., 2023). Furthermore, they can offer immediate feedback helping learners recognize and learn from their mistakes, thus improving their writing accuracy or developing a unique writing style.

LLM chatbots can also be very useful in practicing pronunciation in a foreign language. By integrating text-to-speech technologies, LLMs can produce audio pronunciations of text, generate phonetic transcriptions of words and phrases, helping learners understand the correct pronunciation (Baskara & Mukarto, 2023).

It must also be noticed that LLM Chatbots can promote cultural understanding by providing contextually relevant information about the target language speakers, customs, and traditions and simulate real-world scenarios such as travel, everyday life or professional settings allowing students to practice their language skills more effectively (Baskara & Mukarto, 2023).

It is a common belief among researchers in the field that the use of LLM Chatbots presents a promising potential that should be further explored. However, there are challenges and concerns that need to be addressed. It is obvious that the effectiveness of LLM Chatbots directly depends on the quality of the data they are trained with or have access to. There are, however, other issues arising from their use which concern - or should concern - the academic community. The first of these is the bias in the answers - which is mainly due to the data they have been fed with - and ethical implications (Xu and Yuan, 2021; Rusmiyanto et al., 2023) and privacy concerns, which require responsible use to be ensured. Several other limitations and challenges, such as difficulties in articulating complex or abstract ideas and the risk of producing biased or offensive content are also reported (Baskara & Mukarto, 2023). Another issue that arises is finding a balance between the assistance they can offer and maintaining the ability of users to carry out the necessary tasks themselves without

being solely dependent on them. Specifically for language learning, some researchers express criticism concerning the reliability and the accuracy of these chatbots (Xiao & Zhi, 2023). De la Vall & Araya (2023) express doubts about the ability of chatbots to emulate the cultural and contextual subtleties of language and to comprehend or generate creative or original expressions. Baskara & Mukarto (2023) express skepticism about the ethical issues related to teacher replacement and the impact on the language teaching profession. Finally, as interaction quality greatly impacts students' learning effectiveness, uncertainties regarding this quality can pose challenges in using AI for language learning. To tackle this issue, more advanced AI technology is required and system developers must thoroughly test their designs before deploying new systems (Huang et al., 2023).

7. Language Learning online platforms

The popularity of ChatGPT and the immediate impact it has had on education have led online language course providers to improve their services by incorporating AI technology to offer more personalized learning experiences. Towards this direction, the two major cloud-based language learning platforms, Memrise in 2022 and Duolingo in 2023, introduced their novel LLM-based chatbots and functions (Curry, 2024). MemBot, built on GPT-3, has the capability to customize the learning experience by tailoring the difficulty level for each learner. It continuously assesses the learner's skills and tailors the content to meet individual needs in real-time (Zerkowska, 2024). Similarly, Duolingo Max plan, powered by GPT-4, also uses adaptive algorithms to personalize the learning experience for each user (Bicknell et al., 2023). Both MemBot and Duolingo incorporate a chatbot feature that allows users to practice conversational skills with virtual partners. This is also the case with Microsoft's XiaoICE, which engages users in emotionally intelligent conversations -with a conversational partner-, adapting to their language proficiency and learning style. In addition to utilizing AI algorithms to provide users with personalized language lessons and feedback, these platforms offer interactive lessons, quizzes and exercises to help learners enhance their grammar, vocabulary and speaking skills (Woo & Choi, 2021).

However, the use of LLMs and adaptive learning strategies to maximize the advantages of PLL-based methods extend beyond these two leading platforms. The same methods are also employed by numerous popular online language learning platforms, including Babbel, Mondly, and Rosetta Stone. Babbel offers personalized lesson plans based on user preferences and progress. Its speech recognition technology provides immediate feedback on pronunciation, helping learners improve their speaking skills effectively. MondlyVR offers immersive language learning experiences through virtual reality. Users can practice speaking in simulated real-world scenarios, which helps them develop practical language skills. Rosetta Stone's TruAccent technology uses speech recognition to offer real-time feedback on

pronunciation. The program also adapts to the learner's pace, providing customized lessons that focus on areas needing improvement (Lepcheska, 2023).

Intelligent language tutoring systems, like those mentioned above, are available in mobile apps, online platforms and standalone software. These advancements represent an important shift in language learning and are anticipated to influence future developments in education as well as emerging language learning platforms.

8. Conclusion

Klimova et al. (2023), as well as Shaikh et al. (2023) believe that emerging language learning technologies such as AR/VR based language learning apps, NLP-based language learning platforms, gamification or speech recognition-based language learning, and, of course, LLM chatbots and conversational AI, can drastically change the language learning landscape in the following years. This is an optimistic but also realistic view, resulting from the literature and the very encouraging findings of the research done in the field, which show that AI can be a transformative technology in language education. The findings to date indicate that incorporating AI into language learning environments enhances learning outcomes and provides significant pedagogical benefits by personalizing instruction and fostering learner autonomy. Adaptive learning platforms and AI-based chatbots deliver personalized feedback, practice opportunities and customized learning pathways, resulting in increased engagement and language proficiency.

As AI technologies continue to evolve, it is essential to ensure that these advancements are effectively integrated into educational frameworks. According to Pokrivcakova (2019), current research on AI in language learning has focused on providing personalized learning content, developing tutoring systems, creating VR environments for learners to practice language, translating written or spoken texts, assisting writing, and conducting conversations using chatbots. However, there is a need for further research to identify the best practices for integrating LLMs into education and to address the identified risks.

To start with, it must be thoroughly explored how AI tools can seamlessly be integrated into traditional language teaching methods, and which are the training needs and the most effective ways, strategies and pedagogical approaches and methodologies to blend AI-driven learning with traditional instruction. Improving the quality of results, answers and feedback, adding the possibility to accept multimedia inputs (sound and images), exploring multimodal LLM capabilities (e.g., integrating speech recognition), improving chatbots ability to provide adaptive learning pathways in real-time and their educational effectiveness, are also among the goals that research must focus on. It is also important to address issues surrounding data security, student privacy and the ethical use of AI in education. To this end, ways should be sought to reduce biased information, to exclude the possibility of returning false answers and to enhance cultural sensitivity. Furthermore, to exploit the benefits

of artificial intelligence without risk, it is necessary to take measures and formulate policies or regulations to ensure the safety of students' personal data as well as the reliability of studies.

As emerged from the literature and analyzed in this article, AI applications have the ability to provide adaptive, interactive, and engaging learning experiences capable of meeting the unique needs of each student. The innovative capabilities they offer, combined with human supervision and control as well as critical thinking, can improve language acquisition as well as make learning more enjoyable and accessible for students.

References

[1] Al Jazeera (Sept 28, 2023). Chatgpt can now browse the internet for updated information. Al Jazeera (online). https://www.aljazeera.com/news/2023/9/28/chatgpt-can-now-browsethe-internet-for-updated-information.

[2] Anthropic (2023). Introducing claude. https://www.anthropic.com/news/introducing-claude

[3] Atlas, S. (2023). ChatGPT for higher education and professional development: A guide to conversational AI. https://digitalcommons.uri.edu/cba_facpubs/548

[4] Bai, Y., Jones, A., Ndousse, K., Askell, A., Chen, A., DasSarma, N., ... & Kaplan, J. (2022). Training a helpful and harmless assistant with reinforcement learning from human feedback. arXiv preprint arXiv:2204.05862.

[5] Bang, Y., Cahyawijaya, S., Lee, N., Dai, W., Su, D., Wilie, B., ... & Fung, P. (2023). A multitask, multilingual, multimodal evaluation of chatgpt on reasoning, hallucination, and interactivity. arXiv preprint arXiv:2302.04023.

[6] Baskara, R. & Mukarto, F.X. (2023). Exploring the implications of ChatGPT for language learning in higher education. Indonesian Journal of English Language Teaching and Applied Linguistics, 7(2), 343-358.

[7] Belda-Medina, J., & Calvo-Ferrer, J. R. (2022). Using chatbots as AI conversational partners in language learning. Applied Sciences, 12(17), 8427.

[8] 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), 827-877.

[9] Bicknell, K, Brust, C. & Settles, B. (Feb 4, 2023). How Duolingo's AI learns what you need to learn. IEEE spectrum (online). https://spectrum.ieee.org/duolingo

[10] Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J. D., Dhariwal, P., ... & Amodei, D. (2020). Language models are few-shot learners. Advances in neural information processing systems, 33, 1877-1901.

[11] Caldarini, G., Jaf, S., & McGarry, K. (2022). A literature survey of recent advances in chatbots. Information, 13(1), 41.
[12] Chatlabs. (March 31, 2023). Can Claude.ai Access the Internet? Chatlabs (online). https://writingmate.ai/blog/can-claude-ai-access-the-internet
[13] Chang, Y., Wang, X., Wang, J., Wu, Y., Yang, L., Zhu, K., ... & Xie, X. (2024). A survey on evaluation of large language models. ACM Transactions on Intelligent Systems and Technology, 15(3), 1-45.
[14] Chen, M., Tworek, J., Jun, H., Yuan, Q., Pinto, H. P. D. O., Kaplan, J., ... & Zaremba, W. (2021). Evaluating large language models trained on code. arXiv preprint arXiv:2107.03374.
[15] Chen, X., Meurers, D., & Rebuschat, P. (2022). ICALL offering individually adaptive input: Effects of complex input on L2 development. Language Learning & Technology, 26(1), 1–21. https://hdl.handle.net/10125/73496
[16] Choi, I. C. (2016). Efficacy of an ICALL tutoring system and process-oriented corrective feedback. Computer Assisted Language Learning, 29(2), 334-364. 29:2, 334-364
[17] Chowdhery, A., Narang, S., Devlin, J., Bosma, M., Mishra, G., Roberts, A., ... & Fiedel, N. (2023). Palm: Scaling language modeling with pathways. Journal of Machine Learning Research, 24(240), 1-113.
[18] Chuah, K. M., & Kabilan, M. (2021). Teachers' views on the use of chatbots to support English language teaching in a mobile environment. International Journal of Emerging Technologies in Learning (iJET), 16(20), 223-237.
[19] Colby, K. M., Weber, S., & Hilf, F. D. (1971). Artificial paranoia. Artificial intelligence, 2(1), 1-25.
[20] Condon, S. (Feb 7, 2023). More powerful than ChatGPT': Microsoft unveils new AI-improved Bing and Edge browser Microsoft has tapped OpenAI, the maker of ChatGPT, to help it improve its Bing search engine, while also improving its web browser. ZDNet (online). https://www.zdnet.com/article/microsoft-taps-chatgpt-to-improve-bing-and-the-edge-browser/
[21] Conneau, A., Khandelwal, K., Goyal, N., Chaudhary, V., Wenzek, G., Guzmán, F., ... & Joulin, A. (2020). Unsupervised Cross-lingual Representation Learning at Scale. Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, 8440-8451.
[22] Conway, A. (2023). Microsoft Copilot: What is it, and how does it work? QZDQ (online). https://www.xda-developers.com/bing-chat/
[23] Crider, M. (Feb 8, 2024). Google bard ai is officially gemini, unofficially irrelevant. PCWorld (online). https://www.pcworld.com/article/2230933/google-bard-ai-is-officiallygemini-unofficially-irrelevant.html
[24] Curry, D. (Jan 8, 2024). Language Learning App Revenue and Usage Statistics (2024). Business of Apps (online). https://www.businessofapps.com/data/language-learning-app-market/

[25] Dam, S. K., Hong, C. S., Qiao, Y., & Zhang, C. (2024). A Complete Survey on LLM-based AI Chatbots. arXiv preprint arXiv:2406.16937.
[26] De Barcelos Silva, A., Gomes, M. M., da Costa, C. A., da Rosa Righi, R., Barbosa, J. L. V., Pessin, G. & Federizzi, G. (2020). Intelligent personal assistants: A systematic literature review. Expert Systems with Applications, 147, 113193.
[27] De la Vall, R. R. F., & Araya, F. G. (2023). Exploring the benefits and challenges of AI-language learning tools. International Journal of Social Sciences and Humanities Invention, 10(01), 7569-7576.
[28] Deryugina, O. V. (2010). Chatterbots. Scientific and Technical Information Processing, 37, 143-147.
[29] Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2018). Bert: Pre-training of deep bidirectional transformers for language understanding. arXiv preprint arXiv:1810.04805.
[30] ELMCIP. (2023). Knowledge base (online). https://elmcip.net/creative-work/racter
[31] Fang, Y., Lippert, A., Cai, Z., Chen, S., Frijters, J. C., Greenberg, D., et al. (2022). Patterns of adults with low literacy skills interacting with an intelligent tutoring system. International Journal of Artificial Intelligence in Education, 32(2), 297–322. https://doi.org/10.1007/s40593-021-00266-y.
[32] 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 chat GPT usage. *Computers and Education: Artificial Intelligence*, 100180.
[33] Freeman-Mills, M. (July 24, 2023). What is Bing Chat and how does it work? AI chat explained. Pocket-lint (online). https://www.pocket-lint.com/what-is-bing-chat-explained/
[34] Gillham, J. (January 4, 2024). 75+ claude ai model statistics in q4 2023. Originality AI (online). https://originality.ai/blog/claude-ai-statistics
[35] Haristiani, N. (2019). Artificial Intelligence (AI) chatbot as language learning medium: An inquiry. In Journal of Physics: Conference Series (Vol. 1387, No. 1, p. 012020). IOP Publishing.
[36] Haristiani, N. U. R. I. A., Danuwijaya, A. A., Rifai, M. M., & Sarila, H. (2019). Gengobot: A chatbot-based grammar application on mobile instant messaging as language learning medium. Journal of Engineering Science and Technology, 14(6), 3158-3173.
[37] Hassani, K., Nahvi, A., & Ahmadi, A. (2016). Design and implementation of an intelligent virtual environment for improving speaking and listening skills. Interactive Learning Environments, 24(1), 252-271.
[38] Hetler, A. (October 16, 2023). Bard vs. chatgpt: What's the difference? Techtarget (online). https://www.techtarget.com/whatis/feature/Bard-vs-ChatGPT-Whats-the-difference/
[39] High, R. (2012). The era of cognitive systems: An inside look at IBM Watson and how it works. IBM Corporation, Redbooks, 1, 16.

[40] Huang, T. H., & Wang, L. Z. (2021). Artificial intelligence learning approach through total physical response embodiment teaching on French vocabulary learning retention. Computer Assisted Language Learning, 1. https://doi.org/10.1080/09588221.2021.2008980. –25.
[41] 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.
[42] Huang, X., Zou, D., Cheng, G., Chen, X., & Xie, H. (2023). Trends, research issues and applications of artificial intelligence in language education. Educational Technology & Society, 26(1), 112-131.
[43] Huang, Z. (March 21, 2023). China's first major chatbot doesn't need to be as good as chatgpt. Bloomberg (online). https://www.bloomberg.com/news/newsletters/2023-03-21/baidu-s-ernie-bot-aims-to-be-first-in-chatgpt-free-market-in-china
[44] Jeon, J., Lee, S., & Choe, H. (2023). Beyond ChatGPT: A conceptual framework and systematic review of speech-recognition chatbots for language learning. Computers & Education, 104898.
[45] Jia, J. (2009). CSIEC: A computer assisted English learning chatbot based on textual knowledge and reasoning. Knowledge-based systems, 22(4), 249-255.
[46] Johns, A. (2023). Gpt-4 turbo: Openai announced gpt-4 turbo faster and cheaper. Securecortex Blog (online). https://blog.securecortex.com/2023/11/gpt-4-turbo-openai-announcedgpt-4.html
[47] Kannan, J. & Munday, P. (2018). New trends in second language learning and teaching through the lens of ICT, networked learning, and artificial intelligence. In Fernández Juncal, C. & N. Hernández Muñoz (Eds.). Vías de transformación en la enseñanza de lenguas con mediación tecnológica. Círculo de Lingüística Aplicada a la Comunicación, 76, 13-30
[48] Kasneci, E., Seßler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., ... & Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. Learning and individual differences, 103, 102274.
[49] Kim, H., Yang, H., Shin, D., & Lee, J. H. (2022). Design principles and architecture of a second language learning chatbot. Language Learning & Technology, 26(1), 1–18.
[50] Kim, N.-Y. (2020). Chatbots and Language Learning: Effects of the Use of AI Chatbots for EFL Learning; Eliva Press: Chisinau, Moldova. ISBN 978-1-952751-45-5.
[51] Klimova, B., Pikhart, M., Polakova, P., Cerna, M., Yayilgan, S. Y., & Shaikh, S. (2023). A systematic review on the use of emerging technologies in teaching English as an applied language at the university level. Systems, 11(1), 42.

[52] Lepcheska, T. (2023). What is Babbel and How to Use It? Neural Network Press (online). https://neuralnetworkpress.com/what-is-babbel-and-how-to-use-it

[53] Lewis, M., Liu, Y., Goyal, N., Ghazvininejad, M., Mohamed, A., Levy, O., ... & Zettlemoyer, L. (2020). BART: Denoising Sequence-to-Sequence Pre-training for Natural Language Generation, Translation, and Comprehension. Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, 7871-7880.

[54] Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., Levy, O., Lewis, M., Zettlemoyer, L., & Stoyanov, V. (2019). Roberta: A robustly optimized bert pretraining approach. arXiv. preprint arXiv:1907.11692.

[55] Lozić, E & Štular, B. (2023). ChatGPT v Bard v Bing v Claude 2 v Aria v human-expert. How good are AI chatbots at scientific writing? arXiv preprint arXiv:2309.08636, 2023.

[56] Luo, B., Lau, R. Y., Li, C., & Si, Y. W. (2022). A critical review of state-of-the-art chatbot designs and applications. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 12(1), e1434.

[57] Mageira, K., Pittou, D., Papasalouros, A., Kotis, K., Zangogianni, P., & Daradoumis, A. (2022). Educational AI chatbots for content and language integrated learning. Applied Sciences, 12(7), 3239.

[58] Mavropoulou, E & Arvanitis, P. (2023). Contribution. Implications of Chatbot Integration on Teaching and Learning a Foreign Language. In ICERI2023 Proceedings (pp. 7596-7605). IATED.

[59] Mehdi, Y. (2023). Reinventing search with a new AI-powered Microsoft bing and edge. Microsoft Blog (online). https://blogs.microsoft.com/blog/2023/02/07/reinventing-search-with-a-new-ai-powered-microsoft-bing-and-edge-your-copilot-for-the-web/

[60] Milmo, D. (Nov 5, 2023). Elon musk unveils grok, an ai chatbot with a 'rebellious streak'. The Guardian (online). https://www.theguardian.com/technology/2023/nov/05/elon-musk-unveils-grok-an-ai-chatbot-with-a-rebellious-streak

[61] Min, S., Krishna, K., Lyu, X., Lewis, M., Yih, W. T., Koh, P. W., ... & Hajishirzi, H. (2023). Factscore: Fine-grained atomic evaluation of factual precision in long form text generation. arXiv preprint arXiv:2305.14251.

[62] Mistelbacher, R. (November 1, 2024). The Best AI Assistants Compared: Claude vs Gemini vs ChatGPT vs Mistral vs Perplexity vs CoPilot. Freshvanroot (online). https://freshvanroot.com/blog/best-ai-assistants-compared-2024/

[63] Motlagh, N. Y., Khajavi, M., Sharifi, A., & Ahmadi, M. (2023). The impact of artificial intelligence on the evolution of digital education: A comparative study of openAI text generation tools including ChatGPT, Bing Chat, Bard, and Ernie. arXiv preprint arXiv:2309.02029.

[64] Mubaroq, F. A., Kamalia, K. F., & Zenico, L. (2024). The Advantages and Disadvantages of Using Chat GPT for Writing Skill in Higher School. In *Conference on English Language Teaching* (Vol. 4, No. 1, pp. 178-183).
[65] Obaidoon, S., & Wei, H. (2024). ChatGPT, Bard, Bing Chat, and Claude generate feedback for Chinese as foreign language writing: A comparative case study. Future in Educational Research, 1–21.
[66] OpenAI (2024). GPT-4 Technical Report. (2023). arXiv:cs.CL/2303.08774
[67] Ortiz, S. (February 9, 2024a). What is Google's Gemini AI tool (formerly Bard)? Everything you need to know. ZDNet (online). https://www.zdnet.com/article/what-is-googles-gemini-ai-tool-formerly-bard-everything-you-need-to-know/
[68] Ortiz, S. (May 23, 2024b). What is Copilot (formerly Bing Chat)? Here's everything you need to know. ZDNet (online). https://www.zdnet.com/article/what-is-copilot-formerly-bing-chat-heres-everything-you-need-to-know/
[69] Pichai, S. (2023). Google I/O 2023: Making AI more helpful for everyone [Blog]. Google Blog (online). https://blog.google/technology/ai/google-io-2023-keynote-sundar-pichai/
[70] Pokrivcakova, S. (2019). Preparing teachers for the application of AI-powered technologies in foreign language education. Journal of Language and Cultural Education, 7(3), 135-153.
[71] Radford, A., Wu, J., Child, R., Luan, D., Amodei, D., & Sutskever, I. (2019). Language models are unsupervised multitask learners. OpenAI blog, 1(8), 9.
[72] Raffel, C., Shazeer, N., Roberts, A., Lee, K., Narang, S., Matena, M., Zhou, Y., Li, W., Liu, P. J. (2020). Exploring the limits of transfer learning with a unified text-to-text transformer. Journal of Machine Learning Research, 21(140), 1–67.
[73] Rodriguez, J. (February 28, 2023). Inside sparrow: The foundation of deepmind's chatgpt alternative. Medium (online). https://jrodthoughts.medium.com/inside-sparrow-the-foundationof-deepminds-chatgpt-alternative-854df43569fd
[74] Ruan, S., Willis, A., Xu, Q., Davis, G. M., Jiang, L., Brunskill, E., & Landay, J. A. (2019). Bookbuddy: Turning digital materials into interactive foreign language lessons through a voice chatbot. In Proceedings of the sixth (2019) ACM conference on learning@ scale (pp. 1-4).
[75] Rudolph, J., Tan, S., & Tan, S. (2023). War of the chatbots: Bard, Bing Chat, ChatGPT, Ernie and beyond. The new AI gold rush and its impact on higher education. Journal of Applied Learning and Teaching, 6(1), 364-389.
[76] Rusmiyanto, R., Huriati, N., Fitriani, N., Tyas, N. K., Rofi'i, A., & Sari, M. N. (2023). The role of artificial intelligence (AI) in developing English language learner's communication skills. Journal on Education, 6(1), 750-757.

[77] Scao, T. L., Fan, A., Akiki, C., Pavlick, E., Ili´c, S., Hesslow, D., Castagn´e, R., Luccioni, A. S., Yvon, F., Galle, M. (2022). BLOOM: A 176B-parameter open-access multilingual language model. arXiv. preprint arXiv:2211.05100.

[78] Shaikh, S., Yayilgan, S. Y., Klimova, B., & Pikhart, M. (2023). Assessing the usability of ChatGPT for formal english language learning. European Journal of Investigation in Health, Psychology and Education, 13(9), 1937-1960.

[79] Simplilearn. (April 30, 2024). What is Bing chat? unleash the power of gpt-4 with Bing chat. Simplelearn (online). https://www.simplilearn.com/bing-chat-article

[80] Singh, S., & Thakur, H. K. (2020). Survey of various AI chatbots based on technology used. In 2020 8th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO) (pp. 1074-1079). IEEE.

[81] Sisson, P. (July 25, 2024). In Memphis, an AI Supercomputer from Elon Musk Stirs Hope and Concern. Blooberg (online). https://www.bloomberg.com/news/features/2024-07-25/in-memphis-elon-musk-s-xai-supercomputer-stirs-hope-and-concern

[82] Son, J. B., Ružić, N. K., & Philpott, A. (2023). Artificial intelligence technologies and applications for language learning and teaching. Journal of China Computer-Assisted Language Learning, 2023. De Gryter. https://doi.org/10.1515/jccall-2023-0015

[83] Songsiengchai, S., Sereerat, B. O., & Watananimitgul, W. (2023). Leveraging Artificial Intelligence (AI): Chat GPT for Effective English Language Learning among Thai Students. *English Language Teaching*, *16*(11), 1-68.

[84] Touvron, H., Lavril, T., Izacard, G., Martinet, X., Lachaux, M. A., Lacroix, T., ... & Lample, G. (2023). Llama: Open and efficient foundation language models. arXiv preprint arXiv:2302.13971.

[85] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). Attention is all you need. Advances in neural information processing systems, 30.

[86] Vu, D. C., Lian, A. P., & Siriyothin, P. (2022). Integrating natural language processing and multimedia databases in CALL software: Development and evaluation of an ICALL application for EFL listening comprehension. Computer-Assisted Language Learning Electronic Journal, 23(3), 41-69.

[87] Wallace, R. S. (2009). The anatomy of ALICE (pp. 181-210). Springer Netherlands.

[88] Wang, H. (2024). The Design and Evaluation of Intelligent Computer-Assisted Language Learning Tools for Beginner Learners of Mandarin Chinese (Doctoral dissertation, Trinity College, Dublin).

[89] Wang, S., Christensen, C., Cui, W., Tong, R., Yarnall, L., Shear, L., & Feng, M. (2023). When adaptive learning is effective learning: comparison of an adaptive learning system to teacher-led instruction. Interactive Learning Environments, 31(2), 793-803.

[90] Warren, T. (November 15, 2023). Microsoft rebrands Bing Chat to Copilot, to better compete with ChatGPT / Microsoft has settled on Copilot as its main brand for its AI chatbot, after months of trying to push Bing Chat. The Verge (online). https://www.theverge.com/2023/11/15/23960517/microsoft-copilot-bing-chat-rebranding-chatgpt-ai
[91] Wei, J., Bosma, M., Zhao, V. Y., Guu, K., Yu, A. W., Lester, B., ... & Le, Q. V. (2021). Finetuned language models are zero-shot learners. arXiv preprint arXiv:2109.01652.
[92] Weizenbaum, J. (1966). ELIZA—a computer program for the study of natural language communication between man and Machine. Communications of the ACM, 9(1), 36-45.
[93] Wikipedia (2024). Cleverbot. https://en.wikipedia.org/wiki/Cleverbot
[94] Woo, J. H., & Choi, H. (2021). Systematic review for AI-based language learning tools. arXiv preprint arXiv:2111.04455. https://arxiv.org/pdf/2111.04455
[95] Wu, T., He, S., Liu, J., Sun, S., Liu, K., Han, Q. L., & Tang, Y. (2023). A brief overview of ChatGPT: The history, status quo and potential future development. IEEE/CAA Journal of Automatica Sinica, 10(5), 1122-1136.
[96] Xiao, Y., & Zhi, Y. (2023). An exploratory study of EFL learners' use of ChatGPT for language learning tasks: Experience and perceptions. Languages, 8(3), 212.
[97] Xu, J., Ju, D., Li, M., Boureau, Y. L., Weston, J., & Dinan, E. (2020). Recipes for safety in open-domain chatbots. arXiv preprint arXiv:2010.07079
[98] Xu, X., & Yuan, S. T. (2021). Data privacy protection and research ethics in intelligent learning systems: Current trends and future directions. Journal of Educational Technology & Society, 24(1), 17-32.
[99] Yang, Z. (March 16, 2023). Chinese tech giant baidu just released its answer to chatgpt. MIT Technology Review (online). https://www.technologyreview.com/2023/03/16/1069919/baidu-ernie-bot-chatgpt-launch/
[100] Yang, Z., Dai, Z., Yang, Y., Carbonell, J., Salakhutdinov, R., & Le, Q. V. (2019). XLNet: Generalized Autoregressive Pretraining for Language Understanding. Advances in neural information processing systems, 32. preprint arXiv:1810.04805.
[101] Ye, J., Chen, X., Xu, N., Zu, C., Shao, Z., Liu, S., ... & Huang, X. (2023). A comprehensive capability analysis of gpt-3 and gpt-3.5 series models. arXiv preprint arXiv:2303.10420.
[102] Zakos, J., & Capper, L. (2008). CLIVE–an artificially intelligent chat robot for conversational language practice. In Artificial Intelligence: Theories, Models and Applications: 5th Hellenic Conference on AI, SETN 2008, Syros, Greece, October 2-4, 2008. Proceedings 5 (pp. 437-442). Springer Berlin Heidelberg.
[103] Zerkowska, A. M. (2024). PERSONALIZED LANGUAGE LEARNING IN THE AGE OF AI. Leveraging Large Language Models for Optimal Learning

Outcomes. https://gupea.ub.gu.se/bitstream/handle/2077/81934/Anika%20Zerkowska%20-%20Thesis%20Final.pdf?sequence=1

[104] Zhang, C., Zhang, C., Li, C., Qiao, Y., Zheng, S., Dam, S. K., ... & Hong, C. S. (2023). One small step for generative ai, one giant leap for agi: A complete survey on chatgpt in aigc era. arXiv preprint arXiv:2304.06488.

[105] Zhao, W. X., Zhou, K., Li, J., Tang, T., Wang, X., Hou, Y., ... & Wen, J. R. (2023). A survey of large language models. arXiv preprint arXiv:2303.18223.