www.ijte.net

A Bibliometric Analysis and Systematic Review in AI Chatbots in Language Teaching and Learning

Hui Wen Chua
Universiti Malaysia Kelantan & Universiti Sains Malaysia, Malaysia

Nagaletchimee Annamalai
Universiti Sains Malaysia, Malaysia

To cite this article:

Chua, H.W. & Annamalai, N. (2025). A bibliometric analysis and systematic review in AI chatbots in language teaching and learning. *International Journal of Technology in Education (IJTE), 8*(2), 274-331. https://doi.org/10.46328/ijte.1035

International Journal of Technology in Education
2025, Vol. 8, No. 2, 274-331 https://doi.org/10.46328/ijte.1035

A Bibliometric Analysis and Systematic Review in AI Chatbots in Language Teaching and Learning

Hui Wen Chua, Nagaletchimee Annamalai

Article Info

Article History

Received:
20 October 2024
Accepted:
14 March 2025

Keywords

AI chatbots
Language learning
Language teaching
Bibliometric analysis
Systematic review

Abstract

The role of AI chatbots is undergoing a transformation, where it was firstly used for English native language learning; later, it shifted to the use for learning English as a second language (ESL) and English as a foreign language learning. Lastly, it is used to learn foreign languages. Hence, due to the changes in AI chatbots' role, there is a need for a study to analyse the development of AI chatbots over the years between 2006 and 2024 and their influence on language education. Therefore, bibliometric analysis and systematic analysis study aim to determine the state-of-art topics related to using AI chatbots in language teaching and learning and how different AI chatbots influence the teachers' and students' perspectives on language teaching and learning and students' learning outcomes. The research is concluded as follows: (1) extend the studies toward students/teachers from various regions, language proficiency levels, and communities with different cultural backgrounds, (2) longitudinal research could be employed to see if there is any novelty effect or other changes in the learning outcomes, affective gains and factors influence the use of the AI chatbots over an extended period, (3) focus on developing strategies, language learning model and process, teaching approaches or methods, assistance from teachers and peers and guidelines for integrating AI chatbots, especially with LLMs AI chatbots into curriculum effectively, (4) effects of learning with self-developed AI chatbots or LLMs AI chatbots that are integrated with more intelligence, realistic agents capable of performing several expression, gestures and movements or more additional games, quizzes, and more multimedia elements in enhancing language learning, (5) factors influence teachers and students in acceptance the use of AI chatbots.

Introduction

AI chatbot is a conversational user interface (conversational UI) called a conversational agent, or in other words, a dialogue agent or Artificial intelligence (AI) agent (Divekar et al., 2021). AI chatbot simulates intelligent, human-like conversations, responds to questions, provides answers, and offers synchronous support and tutoring (Kohnke, 2023; Kerly et al., 2007; Pereira & Diaz, 2018). Hence, they build an environment which allows learners to engage in a meaningful verbal exchange with a computer program (Kim et al., 2021; Lee et al., 2020; Wang & Petrina, 2013). The computer program or AI chatbot AI can facilitate audio or text conversations (Shevat, 2017)

to interact with learners in a particular domain or topic by providing intelligent responses in natural language (Abdul-Kader & Woods, 2015; Azwary et al., 2016). The responses given by AI chatbots when they interact with learners are in the form of comments, answers, completing sentences, goal-oriented dialogue, chit-chat dialogue, visual dialogue, or starting new topics (Haristiani & Rifa'i, 2020; Goyal et al., 2018; Huang et al., 2007). Those chatbot activities are in various forms, such as 'text-based, turn-based, task-fulfilling programs, embedded within existing platforms' (Kohnke, 2023; Jain et al., 2018).

Some research proves that AI chatbots can enhance foreign language learning, especially in conversations that involve of speaking skills and listening skills (Ayedoun et al., 2015; Kim et al., 2021; Lin & Mubarok, 2021; Çakmak, 2022; Fathi et al., 2024), language vocabulary (Lee et al., 2023; Ruan et al., 2021; Chen et al., 2020; Rodosthenous et al., 2020; Lee et al., 2023; Polyzi & Moussiades, 2023), writing skills (Escalante et al., 2023; Kwon et al., 2023; Guo et al., 2024); grammar (Kharis et al., 2022), and reading skills (Nozhovnik et al., 2023). However, researchers also report that AI chatbots also have drawbacks. The cons include the decline of learners' interest in the interaction between humans and chatbots over time (Fryer et al., 2017; Thompson et al., 2018; Assim, 2024), unsuitability and irrelevance of chatbots' response in unexpected directions, and the limited ability to generate progressive-mannered questions, which will dampen learners' engagement that may lead to learning frustration (Gallacher et al., 2018; van Wezel et al., 2021; Assim, 2024), less effective toward learners with high-level target language competency compared to learners of low-level competency (Yin & Satar, 2020; Assim, 2024).

AI chatbots underwent a notable advancement in technology in November 2022. ChatGPT, a publicly available cutting-edge Large Language Model (LLM) which can generate human-like conversation through text-to-text or text-to-speech prompts, was developed by OpenAI and was launched (Li et al., 2023). According to Tlili et al. (2023), this breakthrough that comes along with emerging LLMs has revolutionised conventional chatbots and presented remarkable opportunities and challenges to students, teachers, and policymakers, which encourages more research in this field (Yang et al., 2024). Researchers have started integrating ChatGPT and other LLMs into language education to explore its effectiveness. The studies found that the use of ChatGPT and other LLMs could encourage personalised learning, foster complex learning, and even reduce teachers' workload (Farrokhnia et al., 2023; Tlili et al., 2023; Zhai, 2022; Yang et al., 2024). Hence, these studies supported the potential of LLMs as learning tools to provide personalised and interactive simultaneous learning experiences. Therefore, does the emergence of LLM chatbots influence the current development of AI chatbots available in the marketplace? How teachers and students perceive this new technological advancement and its effects on the student's learning outcomes are still questionable.

Literature Review

The Development of AI Chatbots

The research on developing and applying AI chatbots during the COVID-19 pandemic has sharply increased. During that period, AI chatbots played a vital role in the limited language communication during online classes, where they became students' speaking partners for students to practice their language skills without the constraint

of time and space. AI chatbot can develop during the pandemic, it is due to its gradual development since 1956, when the program ELIZA was the first AI chatbot created by Joseph Weizenbaum at MIT (Fryer & Carpenter, 2006; Weizenbaum, 1966), mainly for psychoanalysis purposes. Besides that, the existence of the annual Loebner Prize Competition since 1990, which is a Turing Test (Turing, 1950) to examine the overall ability of chatbots to maintain a coherent conversation, will have an obvious bearing on their potential for use in language learning (Coniam, 2008), makes the foundation of AI chatbots. According to Kim et al. (2022), during that period, the creation of AI chatbots was dedicated to examining the potential of first language (L1) AI chatbots (i.e., those initially built for native speakers of that language), with only a few attempts of design and develop an AI chatbot exclusively intended for second language (L2) learning. The latest development of a cutting-edge Large Language Model (LLM) called ChatGPT was in November 2022 by OpenAI, a Microsoft-sponsored company.

ELIZA was an AI language-oriented program designed to facilitate communication and authentic interaction via text-based input and output using keyword-matching techniques (Kohnke, 2023; Weizenbaum, 1966), especially between a psychoanalyst and a patient by typing in English. The design of the chatbot was based on "Rogerian analysis", where it asked questions based on what the user typed in. The program relied on the keywords or phrases typed by users to provide automatic, programmed responses, mainly on pattern matching. Hence, the software could not understand the conversation (Fryer & Carpenter, 2006; Weizenbaum, 1966). Jabberwack, or another name called Cleverbot, won the Loebner Prize in 2005. It, created by Rollo Carpenter in 1988, takes a notably different approach to other AI chatbots as it tends to learn from every interaction it has with its visitors. Interaction with different users would allow it to learn more than 8 million independently, making it more lifelike. Therefore, Jabberwacky tends to have long conversations with its users, who find it amusing and oddly 'addictive'. Though its responses are often unpredictable or unexpected, this will improve as it continues to learn, and its ability to keep people talking is potentially of significant value for language learning (Fryer & Carpenter, 2006). Thus, the creation of Jabberwack in the first place is "to simulate natural human chat in an interesting, entertaining and humorous manner". However, according to Fryer & Carpenter (2006), through their observations of the patterns of conversational language, the Jabberwacky AI can learn any language with equal ease, which extends its value beyond English as a Foreign Language (EFL) to all Foreign Language Learning (FLL). Hence, Jabberwack has learnt around 30 languages, including Romanized Japanese. The increasing conversation for a dedicated training process between Jabberwack and language teachers could massively improve their abilities regarding where the patterns, such as spelling and grammatical errors, can be taught to respond.

Rapid text and natural language interface development in AI chatbots brings new features. Wallace (1990) created ALICE (Artificial Linguistic Internet Computer Entity) in 1995, one of the largest accessible sources of AI chatbot community on the internet, using an Artificial Intelligence Mark-up Language (AIML), which was the winner of the Loebner Prize in 2000, 2001 and 2004. ALICE-style AI chatbot stores its knowledge of conversation patterns in AIML files. AIML is a derivative of Extensible Mark-up Language (XML). AIML contains objects called AIML objects, which comprise units called topics and categories. The topic is an optional top-level element, which includes a name attribute and a set of categories related to that topic (Shawar & Atwell, 2007; Wang & Petrina, 2013). The basic unit of knowledge in AIML is called a category. There are three types of categories, namely, atomic categories, default categories and recursive categories. Each category is a rule for matching and

converting an input to an output. It consists of a pattern that contains words or sentences provided to the AI chatbot and a template, which is used in matching to find the most appropriate response to users' input and generating the ALICE AI chatbot answer (Shawar & Atwell, 2007; Wang & Petrina, 2013). Another AI chatbot that adopts AIML files is Kuki, previously known as Mitsuku, developed by Steve Worswick. It can stimulate the most human-like conversation in the world by providing human-like responses to users' input and even understanding mood in users' typed language. Hence, it has been a five-time Loebner Prize winner from 2013 to 2019 (Yang et al., 2022).

The other offshoot of Wallace's A.L.I.C.E.– March 2002" ALICE artificial intelligence program, called Lucy, hosted on Pandorabots website. The design of Lucy is more "language tutor" than ALICE. She is trained based on the commercial AI chatbot Lucy's World, which focuses on different topics, including helping visitors, hotel English, giving directions, English for travelling, and restaurant English. Lucy offers users over 1000 sentences on a specific subject in each world. The design of Lucy as an online language robot is to help English 101 learners review English grammar and vocabulary learned from Lucy's world. Hence, learners can converse extensively with Lucy by speaking to their computer through the microphone. With an advanced speech recognition system, Lucy can provide learners feedback on their pronunciation and guide them through valuable exercises to improve their pronunciation and accuracy. (Wang & Petrina, 2013).

The computer Simulation in Educational Communication (CSIEC) system developed by Jia (2009) in China exceeded the old ELIZA-like keyword-matching mechanism. The creation of the CSIEC system was based on logical reasoning and inference directly through syntactical and semantic analysis of textual knowledge. The AI chatbot applied the Natural Language Markup Language (NLML) approach to generate communicative responses. The statistical analysis results of users' behaviour in using CSIEC show that users prefer free chatting without spelling and grammar checking.

Mondly is a platform that has been available since 2013 for learning 33 languages in iOS, Android, and web systems. Mondly AI chatbot provides adaptive lessons that encourage users to practice what they learn in everyday situations, such as making food and drinks, ordering in restaurants, making restaurant reservations, etc. Mondly, the AI chatbot strives to be the closest to real interaction as it understands spoken language, can respond with a human voice, changes outfits based on the discussion topic, and uses gestures and facial expressions to create dynamic dialogues. AI chatbot technology aims to promote the most realistic communication possible, keeping memories and thoughts and speaking just like a human, the ability to be a friend and someone who will help, teach and maintain an emotional connection. Hence, Mondly creates specific learning patterns or personalised scripts for all its languages for groups of learners capable of combining certain traits, such as similar mistakes (Poseletska et al., 2023). Moreover, Mondly is also famous as it employs flashcards to help users memorise words and pronunciation (Kohnke, 2023).

Google's Dialogflow is a scenario-based AI chatbot builder, which was known initially as Speaktoit in 2010, founded by Gelfenbeyn, Artem Goncharuk, and Leonid Zolotarev, before it transitioned to Api.ai in 2014, and in 2017 it is rebranding to Google's Dialogflow. Google's Dialogflow is a natural language processing (NLP)

platform that enables developers to create conversational user interfaces for applications, devices, and websites, such as AI chatbots and voice assistants. Dialogflow interprets user input, analyses it using machine learning, and generates appropriate responses. Its advantage is that it is easy to develop an AI chatbot by typing each conversion into an input field called 'Intent'. In addition, it instantly generates the Uniform Resource Locator, which provides an environment like practising conversations on asocial networking service. In the accessible version of DialogflowTM, the learner's input can be voice or text, but the AI chatbot's utterances are only in text. Also, even if the learner speaks in voice on the dialogue screen, it is presented as text through speech recognition (Shin et al., 2024). Everyone without any programming background can build their AI chatbots with Google's Dialogflow, and it is easy to integrate with various communication platforms like Google Assistant, Facebook Messenger, Slack, Telegram, and many others. Easy access to Google's Dialogflow enables developers to build conversational interfaces that can be deployed seamlessly across multiple platforms.

Within the last two years, there have been emerging highly sophisticated generative pre-trained transformer (GPT) large language models (LLMs), such as ChatGPT and Bing Chat from OpenAI's Generative Pre-Trained Transformer (GPT) family, as well as Google Bard from the LaMDA family (Escalante et al., 2023; Liu et al., 2024), which can generate human-like conversation through text-to-text or text-to-speech prompts (Li et al., 2023). These AI technologies operating through pre-trained large language models (LLMs) built on extensive web corpus data often take the form of AI chatbots. These chatbots have demonstrated remarkable potential in accomplishing natural-language tasks with linguistic accuracy and generating contextually appropriate texts that recognise language varieties, cultural references, and genres (Tam, 2023; Liu et al., 2024). Hence, all users use these LLMs to send questions in the prompt. Therefore, how do these LLMs have implications for language teaching and learning? The teachers' and students' perspectives are essential, especially regarding factors influencing their usage of these new applications.

The Taxonomy of AI Chatbot Models

The development of AI chatbots over the years has led Kojouharov (2016) to create a taxonomy of the chatbot model in Figure 1 below to illustrate the challenges of achieving free-flowing and accurate responses (Kim et al., 2022). The figure shows that AI chatbots can divide into open—or closed-domain conversations, retrieval-based, and generative-based responses.

The open-domain conversation sets no direction like human-to-human conversational interaction, where it expects users to engage in extended discussions on a wide range of topics. Many language data and sophisticated learning algorithms related to diverse issues and world knowledge are required to make the task possible. Hence, the open-domain AI chatbots in the market include Jabberwack and ChatGPT. Meanwhile, for the closed domain, a conversation is more goal-oriented and functional, with domain knowledge, topics and contexts for conversation being predefined. Most of this domain, especially conversational user interface (UI), is designed to fulfil the program tasks as efficiently as possible. Therefore, the expected dialogue responses could be limited to pre-design functions and topics, and the closed-domain conversation AI chatbots in the market are Mondly and ALICE. According to Kim et al. (2022), due to the predesigned tasks and topics of closed-domain AI chatbots, their

recognition rate is expected to be higher than in open-domain AI chatbots under the same conditions.

On the other hand, one of the natures of AI chatbot responses is retrieval-based responses, which respond to users through scripted rules. The responses are generated through rule-based intent matching or machine learning training (Kim et al., 2022). The AI chatbots in the market mostly use retrieval-based responses, such as Jabberwack, ALICE, Mondly, etc. Conversely, a generative-based AI chatbot produces new responses from scratch. Generative models are typically based on deep learning and machine translation techniques, which are difficult to achieve. The current AI chatbots with generative-based characteristics are ChatGPT, Bing Chat, etc. The recent development of big language models such as GPT-4, especially ChatGPT 4 and Switch C (Fedus et al., 2021) has made a big wave in society can advance the realisation of generative AI, with its promise of open domain functionality. However, these models' language output has not yet been done when applied to developing AI chatbots for language learners. Moreover, researchers also argue their worries that these models may "overrepresent hegemonic viewpoints and encode biases potentially damaging to marginalised populations" (Bender et al., 2021, p. 610; Godwin-Jones, 2021), which need to be carefully monitored when developing the AI chatbots for language learners.

As we can see, the development of AI chatbots takes a long time, from psychoanalysis purposes to language learning. The development of AI chatbots started with the reliance on the keywords or phrases typed by users to provide automatic, programmed responses, until its remarkable demonstration of remarkable potential in accomplishing natural-language tasks with linguistic accuracy and generating contextually appropriate texts that recognise language varieties, cultural references, and genres. Therefore, this study aims to discover the state of the art of AI chatbots in language teaching and learning when new language models emerge and their implications in language teaching and learning.

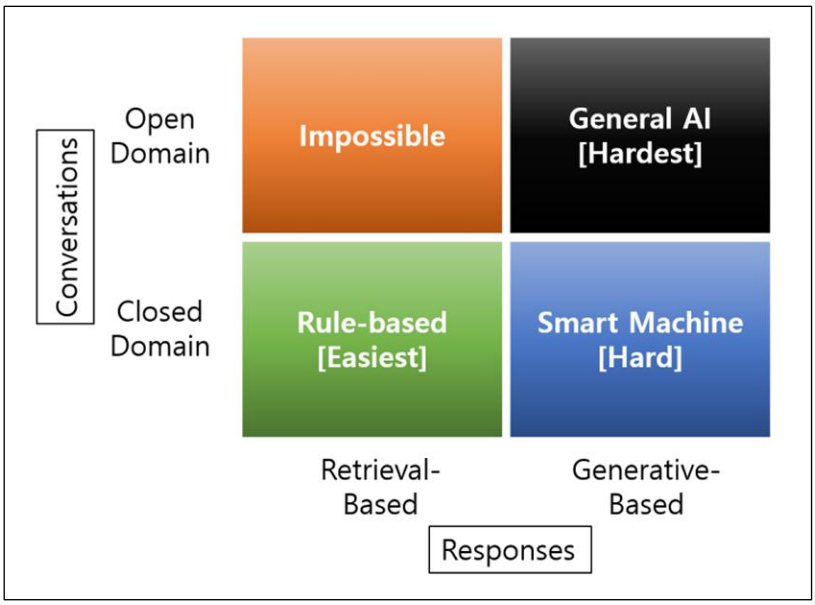


Figure 1. Taxonomy of an AI Chatbot Model (adapted from Kojouharov, 2016)

Since the development of AI chatbots has been around for a certain period, we would like to find out the

backgrounds and trends of AI chatbot development, how different types of chatbots influence the teachers' and students' perspectives on language teaching and learning and students' learning outcomes. Thus, we set up the research questions as below:

1. What are state-of-art topics related to using AI chatbots in language teaching and learning from 2006 to 2024?
2. What are the top ten countries, authors, organisations and publications among the studies on AI chatbot applications in language teaching and learning?
3. How do different AI chatbots influence the teachers' and students' perspectives on language teaching and learning and students' learning outcomes?

Methods

The authors applied bibliometric analysis to respond to the first and second research questions related to the state-of-art topic of using AI chatbots in language teaching and learning. The analysis is run by the graphical user interface-based software of VOSviewer (Van Eck & Waltman, 2010) to help in generating tables, networks and maps representing the results obtained by the techniques, such as co-citation analysis, bibliographic coupling, co-authorship analysis and co-occurrence of keywords (Donthu et al., 2021; Jaleniauskienė et al., 2023). Meanwhile, the authors apply a systematic review, PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses), to answer the first and third research questions. We abstracted the data from the Web of Science (WOS) and Scopus database. The word sequences: "chatbot" OR "chatbots" OR "AI chatbots" OR "interactive personal assistant" AND "language learning" OR "language education" OR "language teaching" OR "second language learning" were used for the search. The authors did not set a year limit, so the authors recovered all searches until April 27, 2024, to find out when the publication research on the application of AI chatbots in language education.

The authors identified 330 articles, 165 from the WOS and 215 from the Scopus databases, as shown in Figure 2. Later, we removed 103 duplicated articles and 11 proceedings and conferences that contained more than one article. The remaining 216 articles have proceeded to be screened by title and abstract. Next, we excluded articles if (1) an AI chatbot was applied but was not related to language education, (2) an article in the form of a review, (3) an abstract was not available and (4) an editorial paper. Therefore, we only included 170 articles and excluded 46 articles. The remaining articles underwent full-text screening. The articles were excluded if (1) the article was in the form of a review, (2) only the creation process of an AI chatbot for language learning, (3) not empirical, (4) overview of AI chatbots, (5) article not written in English, (6) full article not available. Finally, only 86 articles were considered for synthesis. During the selection process, two reviewers performed the literature filtering. If the two could not agree on the selection, a third rater was consulted to make a final decision. Inter-rater reliability reached a satisfactory level (k = 0.729).

Table 1. Inter-rater Reliability

Symmetric Measures		Value	Asymp. Std. Error[a]	Approx. T[b]	Approx. Sig.
Measure of Agreement	Kappa	.729	.052	9.559	.000
N of Valid Cases		170			

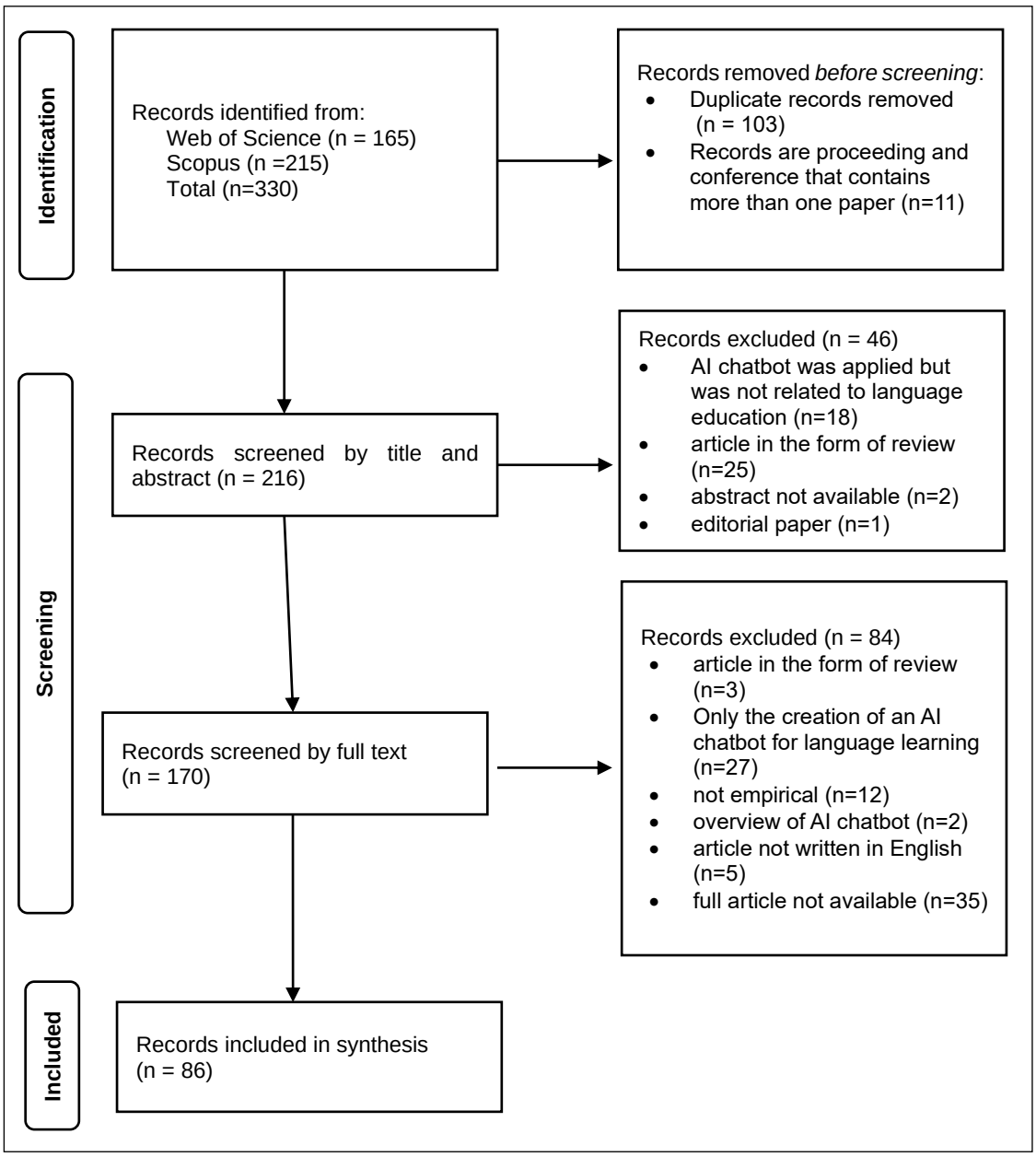


Figure 2. Flowchart of Study Selection Process

Results

RQ1: What are state-of-art topics related to using AI chatbots in language teaching and learning from 2006 to 2024?

Cluster mapping based on keywords. To visualise the state-of-art topic related to the study and select which aspects are suitable for this research. The authors apply the *co-occurrence of all keywords* function of VOSviewer to create a map based on the bibliographic data obtained from WOS and Scopus and select *co-occurrence* as the type of analysis, *full counting* as counting method, and *all keywords* as a unit of analysis. A list of thesauruses in Table 1 is uploaded to VOSviewer to eliminate or consolidate the terms before the following procedure. The authors set the minimum number of occurrences of a keyword at 3. Of 955 keywords, 78 keywords met the

threshold. For each of the 78 keywords, the total strength of the co-occurrence links with other keywords was calculated. The keywords with the greatest total link strength were selected, and the final keyword network is shown in Figure 1. The keywords are categorised into eight clusters, shown in Table 2 and Figure 3, where links, total link strength, occurrence, and average publication year are provided. According to the list, the authors select ten terms with the highest total link strength: chatbot (Occurrences=104, Total link strength=388), language learning (Occurrence= 62, Total link strength=259), artificial intelligence (ai) (Occurrence=55, Total link strength=179), learning systems (Occurrence=23, Total link strength=153), student (Occurrence=23, Total link strength=132), computer-aided instruction (Occurrence=15, Total link strength=104), e-learning (Occurrence=18, Total link strength=103), ChatGPT (Occurrence=29, Total link strength=97), natural language processing systems (Occurrence=9, Total link strength=59) and computer-assisted language learning (call) (Occurrence=13, Total link strength=55).

Table 2. A Thesaurus for VOSviewer to Eliminate or Consolidate Terms

Label	**Replace by**
ai	artificial intelligence (ai)
ai chatbot	artificial intelligence chatbot
artificial intelligence	artificial intelligence (ai)
call	computer-assisted language learning (call)
chatbots	chatbot
computer assisted language learning	computer-assisted language learning (call)
computer-assisted language learning	computer-assisted language learning (call)
conversational agents	conversational agent
educational chatbots	educational chatbot
efl	english as a foreign language
english	english language
english languages	english language
english learning	english language learning
english teaching	english language teaching
english-as-a-foreign-language	english as a foreign language
errors correction	error correction
esl	english as a second language
human-computer interaction	human computer interaction
humans	human
l2	second language
l2 writing	second language writing
machine-learning	machine learning
nlp	natural language processing
pattern-matching	pattern matching
students	student

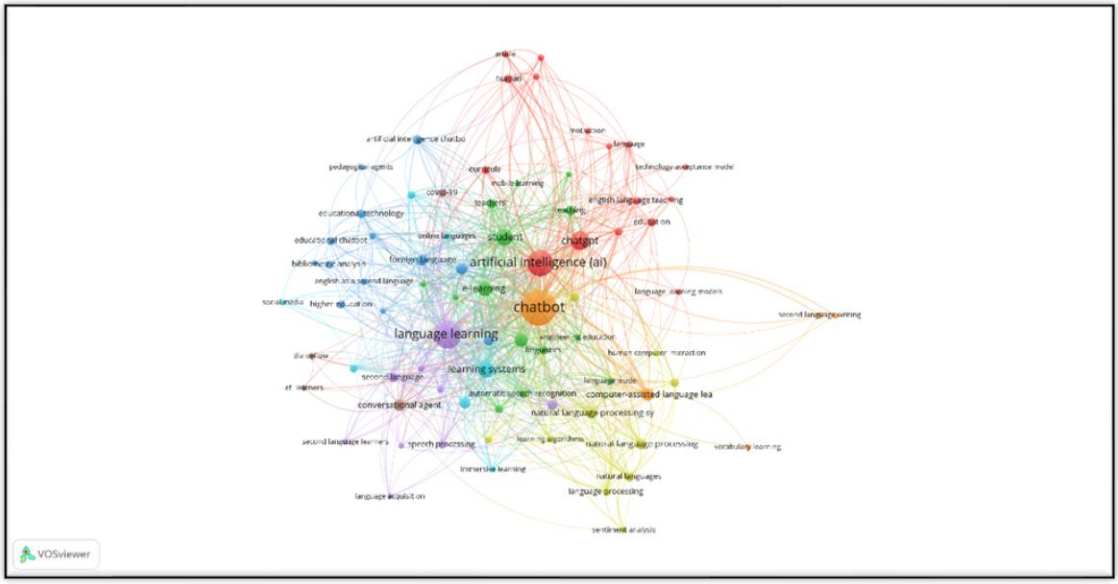


Figure 3. Clustering Map of Keywords

Table 3. Items Categorised according to Cluster with Links, Occurrence, Total Link Strength and Average Year of Publication

Cluster	Item	Links	Total link strength	Occurrence	Average year of publication
1	article	11	23	4	2022.00
	artificial intelligence (ai)	57	179	55	2022.49
	chatgpt	41	97	29	2023.38
	controlled study	12	24	4	2021.50
	curricula	19	29	5	2021.20
	education	18	24	6	2019.67
	english language learning	23	29	5	2019.80
	english language teaching	19	30	5	2023.00
	human	14	32	6	2021.17
	language	18	25	4	2022.50
	language development	12	26	4	2020.50
	language learning model	7	8	3	2023.33
	learning	10	14	3	2022.67
	motivation	5	6	3	2020.33
	technology	9	11	4	2021.75
	technology acceptance	3	4	3	2024.00
2	automatic speech recognition	20	36	5	2022.20
	computer aided instruction	42	104	15	2020.93
	e-learning	40	103	18	2021.17

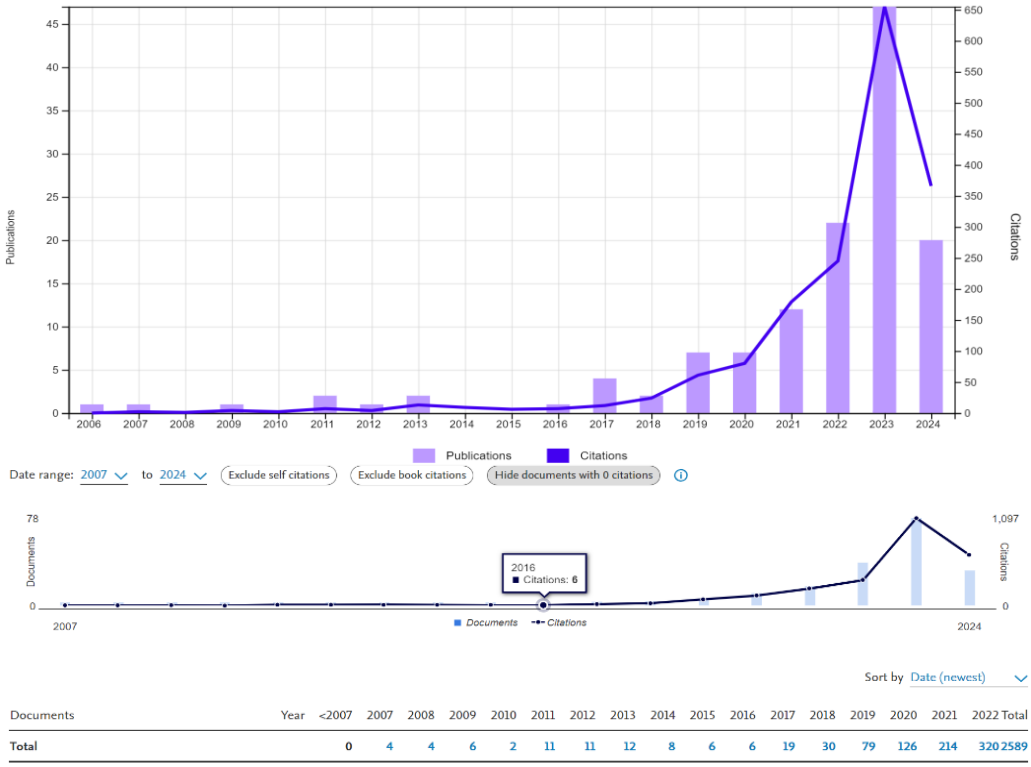
Cluster	Item	Links	Total link strength	Occurrence	Average year of publication
	engineering education	22	28	3	2023.33
	language model	20	23	3	2023.33
	linguistic	24	42	6	2021.17
	mobile learning	9	11	1	2020.33
	natural language understanding	13	17	3	2021.33
	personal assistants	17	20	3	2019.33
	speech recognition	24	40	5	2022.20
	student	49	132	23	2021.87
	teachers'	26	51	7	2022.00
	teaching	24	52	8	2021.88
	teaching and learning	15	18	3	2022.00
3	artificial intelligence chatbot	17	21	7	2022.57
	bibliometric analysis	8	9	5	2023.40
	digital learning	7	7	3	2022.00
	educational chatbot	17	19	5	2021.40
	educational technology	20	26	6	2021.67
	english as a second language	12	16	4	2018.50
	english language	34	55	11	2022.18
	foreign language	33	55	9	2020.44
	foreign language learning	16	22	4	2020.75
	higher education	8	9	5	2021.40
	language education	22	38	8	2022.62
	pedagogical agents	5	6	3	2021.67
	user interfaces	12	14	3	2020.67
4	human computer interaction	10	13	3	2016.67
	language processing	19	38	5	2022.20
	learn+	24	39	7	2022.43
	learning algorithms	17	26	4	2022.25
	machine learning	15	25	6	2022.33
	natural language processing	24	51	9	2020.67
	natural language processing systems	31	59	9	2020.11
	natural languages	20	43	7	2022.00
	sentimental analysis	7	12	3	2021.33
	virtual reality	17	23	4	2019.75
5	computational linguistics	26	44	8	2020.25
	deep learning	22	27	4	2020.75
	dialogue systems	15	23	4	2021.50

Cluster	Item	Links	Total link strength	Occurrence	Average year of publication
	human-machine	9	12	3	2017.67
	language acquisition	7	10	3	2021.33
	language learning	63	259	62	2021.58
	second language	27	41	8	2021.00
	second language learners	7	12	3	2021.33
	speech processing	20	31	5	2021.40
6	english as a foreign language	18	37	11	2022.27
	immersive learning	14	22	3	2021.00
	learning experiences	15	23	5	2022.80
	learning systems	49	153	23	2022.00
	online languages	14	21	3	2022.33
	online learning	16	29	5	2022.40
	social media	6	9	3	2021.67
7	chatbot	75	388	104	2021.52
	computer assisted language learning (call)	28	55	13	2019.69
	second language writing	5	9	4	2023.25
	technology-enhanced language learning	3	6	3	2022.67
	vocabulary learning	2	3	3	2022.33
8	conversational agent	26	53	12	2021.17
	covid-19	14	21	5	2022.00
	dialogflow	7	10	3	2023.00
	efl learners	6	8	3	2023.00

Cluster mapping based on published and cited times. The published and cited times concerning AI chatbots and language education since 2006 for WOS and 2007 for Scopus until 2024 are shown in Figure 4 for both WOS (upper) and Scopus (lower) databases. According to the WOS database, the number of publications regarding AI chatbots and language education was scarce; there were not more than five publications annually until 2020. There were no publications regarding AI chatbots and language education during that period, especially in 2008, 2010, 2014 and 2015. However, the number of publications in 2021 had reached 13, bursting until more than 45 in 2023. The number of publications for the year 2024 remained increasing as it reached 20 at the end of the first quartile of the year. The number of citations slightly fluctuated between 2006 and 2017, at around 25. The number of citations increased gradually until 80 in 2020 and remained increased to 250 in 2022. The number of citations increased drastically to 650 in the year 2023. In the year 2024, the number of citations reached 370 in the first quartile of the year.

Meanwhile, for the Scopus database, publications fluctuated between 2009 to 2011 and 2013 to 2017. After 2016,

the number of publications increased drastically from 2017 until 2022. There was a low number of citations between 2007 and 2016, with only six. The number of citations started to increase gradually between 2017 and 2022, and the numbers increased significantly in 2023 as it reached 1097 citations. The increase in the number of publications regarding AI chatbots and language education in 2023 may have been caused by the emergence of ChatGPT in November 2022, which influenced more research into the application of ChatGPT in language education.



Documents	Year	<2007	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	Total
Total		0	4	4	6	2	11	11	12	8	6	6	19	30	79	126	214	320	2589

Figure 4.Times Cited and Publication (WOS upper, Scopus lower)

Based on Figure 5, the systematic review of the data also supports the result of Cluster mapping based on published and cited times in WOS and Scopus, where AI chatbots have undergone considerable development during the year 2020 and onwards. The number of AI chatbot studies before 2020 was very scarce, with only one research for self-developed AI chatbots without integrating social media applications or platforms and three research studies using more than one AI chatbot application between 2007 and 2015. From 2016 until 2019, five studies were conducted on self-developed AI chatbots without social media applications or platform integration, and three were conducted using only one AI chatbot application. However, between 2020 and 2022, there was a spike increase in self-developed AI chatbots without integrating social media applications or platforms and self-developed AI chatbots with integrating social media applications or platforms, which is nine and 11 research. Most category AI chatbots underwent increments except for only one AI chatbot, which declined from three to two research studies. The number of studies for self-developed AI chatbots without integrating social media applications and using more than one AI chatbot increased in 2023 and onward, with 12 and six research studies. However, integrating self-developed AI chatbots with social media applications or platforms seems to have declined in three research studies. In 2023 onward, there will be more research on the application of Large

Language Model (LLM) or Generative AI (GenAI) AI chatbot applications, such as ChatGPT and Bing Chat.

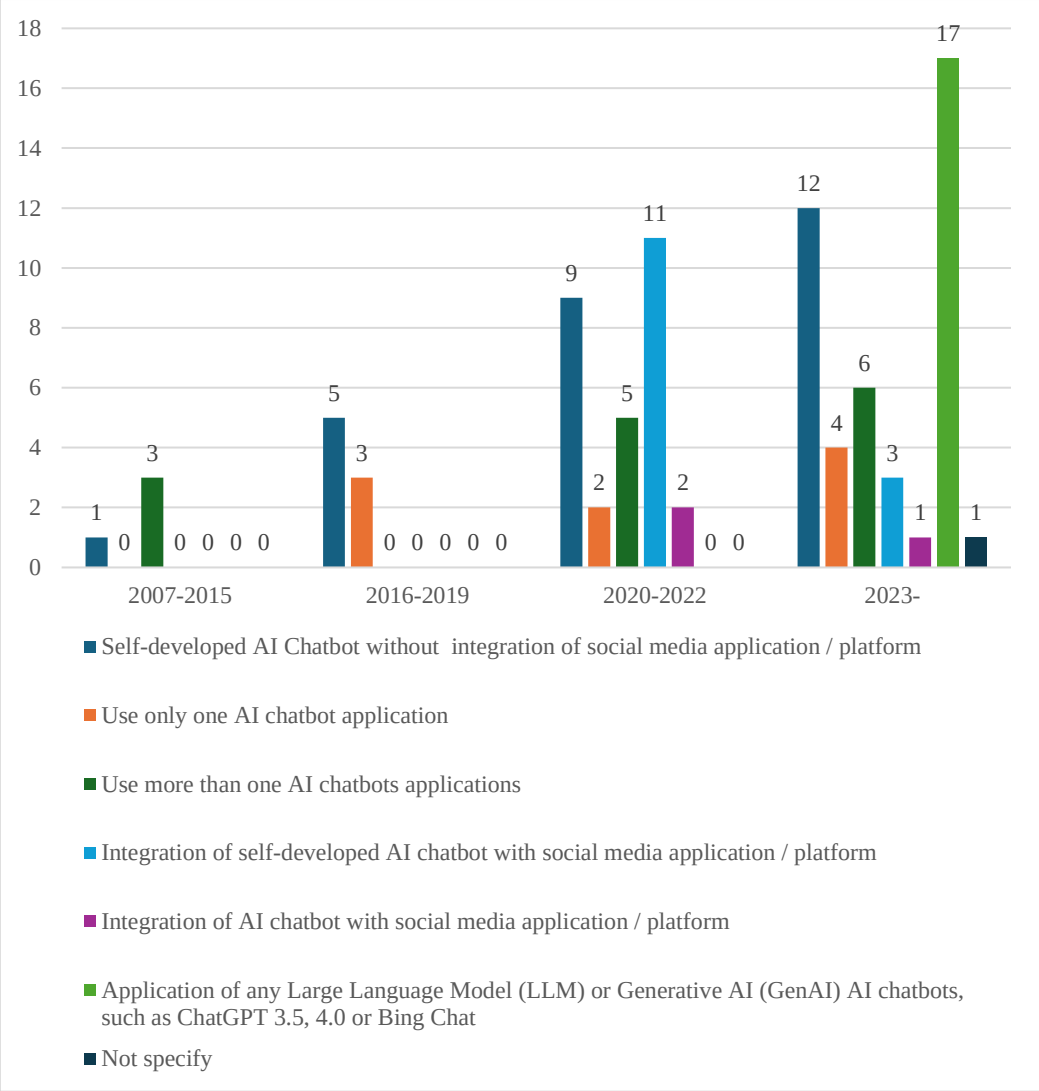


Figure 5. The Trends of AI Chatbots between the Year 2007 and 2024

Based on the data given, the authors found that the developments of chatbots in language learning are closely related to the emergence of technology in artificial intelligence (AI). The maturing growth of AI is influenced by the progress of natural language processing systems in language systems. Hence, these developments also enhanced the expansion of computer-assisted language learning (CALL), e-learning and computer-aided instruction. The maturity of ChatGPT, a type of AI technology, pushes the development of AI technology to its peak. Therefore, the data given by Vosviewer in the bibliometric analysis derives the third research question, which is about the developments of learning systems in AI chatbots and how they influence teachers' language teaching and students' language learning.

Moreover, the average year of publication for each item in Figure 6 will project future suggestions for AI chatbot applications in language teaching and learning. The top ten average years of publication for 2023 are ChatGPT, English language teaching, language learning model, technology acceptance, engineering education, language

model, bibliometric analysis, second language writing, dialogflow and EFL learners. Thus, these will provide some information regarding future study of this research.

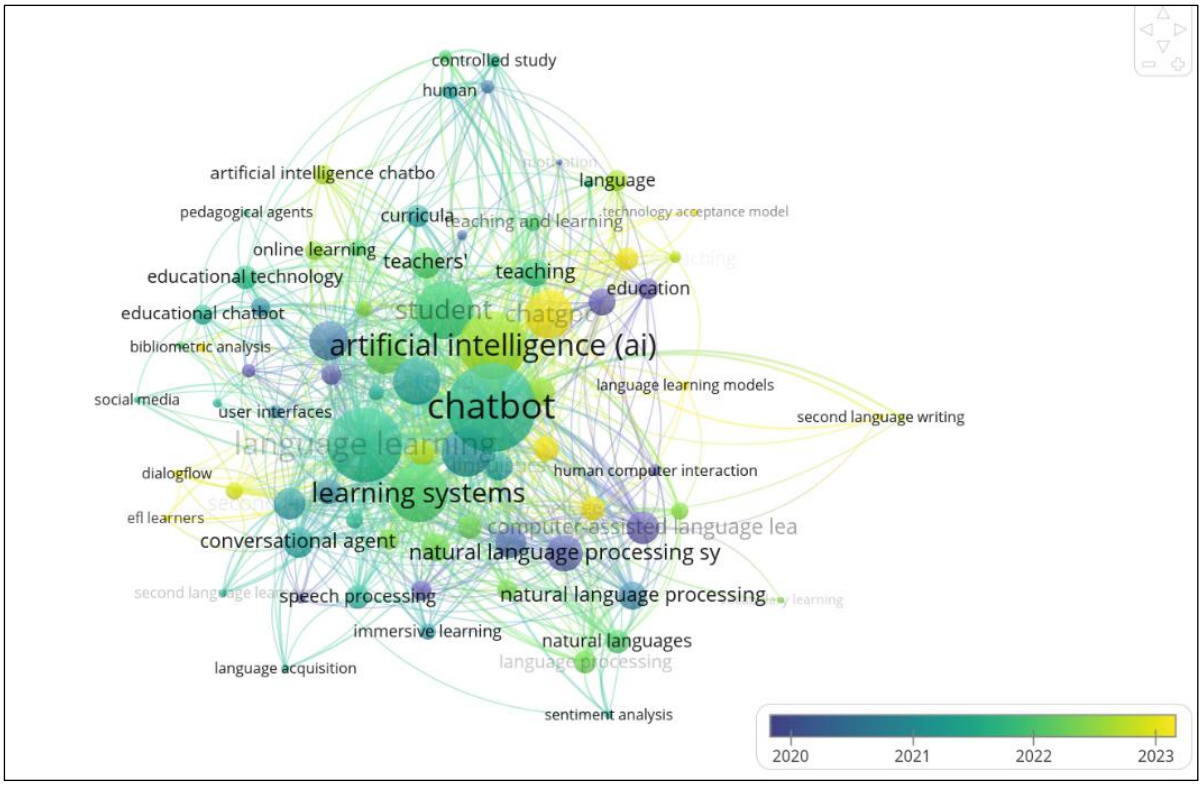


Figure 6. Average Year of Publication

R2: What are the top ten countries, authors, organisations and publications among the studies on AI chatbot applications in language teaching and learning?

The authors analyse the co-occurrence to identify the emerging issues in the current studies on AI chatbot applications in language education. It is essential to undergo a co-citation analysis of the authors to identify their contributions to the field by looking at the citations in their research. Apart from that, the organisations and countries active in the field must also be identified. Hence, the authors run a co-citation analysis onto the authors. Meanwhile, a co-authorship analysis was run to identify countries and organisations actively involved in research regarding AI chatbots in language education.

In the co-citation analysis, the authors select 20 minimum citations per author and discover that 49 authors meet the threshold. The top ten authors are listed with the number of citations and total link strength in Table 4 and the clustering map in Figure 7. The Top ten authors are selected based on their total link strength: Fryer L. K. (Citations: 218; Total link strength: 4155), Thompsom A. (Citations: 84; Total link strength: 1830), Carpenter R. (Citations: 76; Total link strength: 1565), Coniam D. (Citations: 62; Total link strength: 1384), Jeon J. (Citations: 46; Total link strength: 1331), Kim N.-Y. (Citations: 63; Total link strength: 1322), Huang W. (Citations: 60; Total link strength: 1186), Hew K.F. (Citations: 57; Total link strength: 1177), Zou D. (Citations: 56; Total link strength: 1040) and Lee S. (Citations: 50; Total link strength: 1002).

Table 4. Top Ten Most Cited Author

No.	Author	Citations	Total link strength
1	Fryer L. K.	218	4155
2	Thompsom A.	84	1830
3	Carpenter R.	76	1565
4	Coniam D.	62	1384
5	Jeon J.	46	1331
6	Kim N.-Y.	63	1322
7	Huang W.	60	1186
8	Hew K.F.	57	1177
9	Zou D.	56	1040
10	Lee S.	50	1002

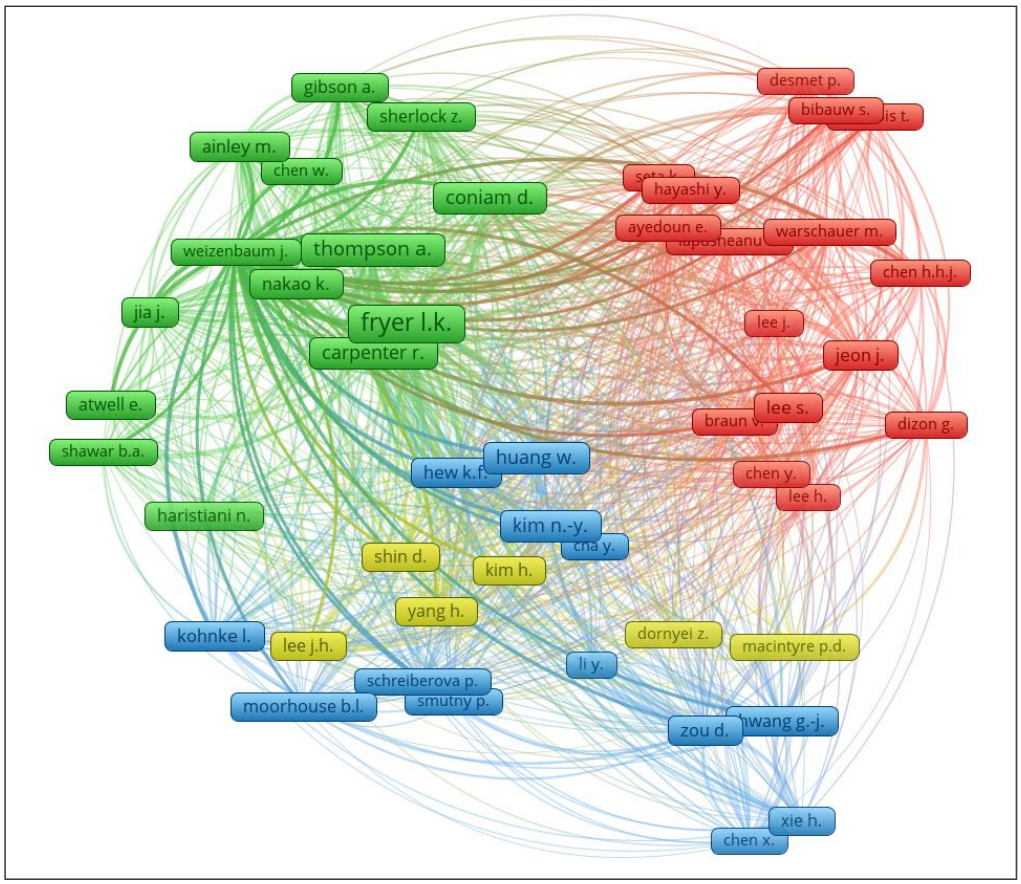


Figure 7. Most Cited Author

Besides that, the authors run the co-authorship analysis for countries actively involved in the AI chatbot application for language education. The authors choose two minimum number of documents of a country and five minimum citations of a country. The results show that 57 countries and 33 countries have met the thresholds. The top ten countries are selected based on their total link strength are: China (Documents: 27; Citations: 351; Total link strength: 19), Hong Kong (Documents: 26; Citations: 846; Total link strength: 16), United Kingdom

(Documents:12; Citations: 52; Total link strength: 15), United States (Documents: 31; Citations: 218; Total link strength: 14), South Korea (Documents: 17; Citations: 177; Total link strength: 10), Australia (Documents: 9; Citations: 230; Total link strength: 8), Italy (Documents: 9; Citations: 15; Total link strength: 7), Spain (Documents: 6; Citations: 79; Total link strength: 6), Ireland (Documents: 3; Citations: 26; Total link strength: 6) and Cyprus (Documents: 2; Citations: 10; Total link strength: 6), as shown in Table 5and Figure 8.

Table 5. Top Ten Country

No.	Country	Documents	Citations	Total link strength
1	China	27	351	19
2	Hong Kong	26	846	16
3	United Kingdom	12	52	15
4	United States	31	218	14
5	South Korea	17	177	10
6	Australia	9	230	8
7	Italy	4	15	7
8	Spain	6	79	6
9	Ireland	3	26	6
10	Cyprus	2	10	6

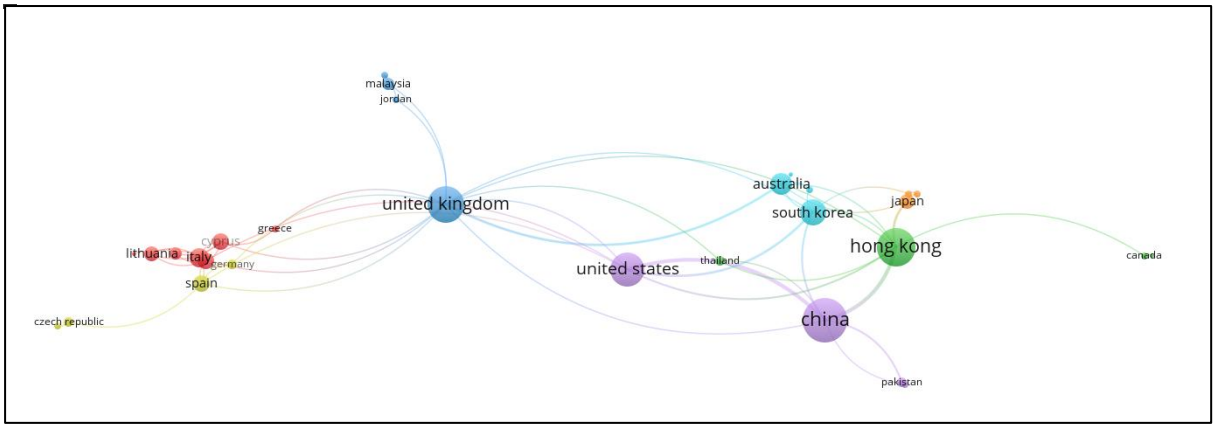


Figure 8.Countries with Most Citation

The authors also run co-authorship analyses toward the organisations involved in AI chatbot applications in language education. The authors select the organisations with at least two documents and have the least number of four citations. Hence, the top ten organisations that fulfil the minimum requirements with the highest total link strength: Columbia University, United States (Documents: 2; Citations: 5; Total link strength: 3), University of California, Davis, United States (Documents: 2; Citations: 5; Total link strength: 3), Chung-Ang University, South Korea (Documents: 2; Citations: 28; Total link strength: 2), Gwangju National University of Education, South Korea (Documents: 2; Citations: 28; Total link strength: 2), The University of Hong Kong, Hong Kong (Documents: 6; Citations: 358; Total link strength: 2), Indiana University, Bloomington, In, United States (Documents: 2; Citations: 15; Total link strength: 0), National Taiwan University of Science and Technology,

Taiwan (Documents: 2; Citations: 116; Total link strength: 0), Kaunas University of Technology, Lithuania (Documents: 2; Citations: 8; Total link strength: 0), Service de Chirurgie Orthopédique, Clinique Aguiléra, 21, Rue De L'estagnas, Biarritz, France (Documents: 2; Citations: 4; Total link strength: 0) and Stanford University, Stanford, CA, United States (Documents: 2; Citations: 25; Total link strength: 0) as shown in Table 6.

Table 6. Top Ten Organisation

No.	Organisation	Documents	Citations	Total link strength
1	Columbia University, United States	2	5	3
2	University of California, Davis, United States	2	5	3
3	Chung-Ang University, South Korea	2	28	2
4	Gwangju National University of Education, South Korea	2	28	2
5	The University of Hong Kong, Hong Kong	6	358	2
6	Indiana University, Bloomington, In, United States	2	15	0
7	National Taiwan University of Science and Technology, Taiwan	2	116	0
8	Kaunas University of Technology, Lithuania	2	8	0
9	Service de Chirurgie Orthopédique, Clinique Aguiléra, 21, Rue De L'estagnas, Biarritz, France	2	4	0
10	Stanford University, Stanford, CA, United States	2	25	0

RQ3: How do different AI chatbots influence the teachers' and students' perspectives on language teaching and learning and students' learning outcomes?

The different types of AI chatbots can be divided into LLM or Gen AI chatbots like ChatGPT 3.5, 4.0 or Bing Chat, self-developed AI chatbots with and without integration of social media application or platform, use only one AI chatbot and use of more than one AI chatbots that available in the market, and integration of AI chatbot available in the market with social media application or platform. There are 27 research studies on self-developed AI chatbots without integration of social media applications or platforms, 17 research studies in LLM or Gen AI chatbots like ChatGPT 3.5, 4.0 or Bing Chat, 14 research study in self-developed AI chatbot with integration of social media application or platform, 12 research on using more than one AI chatbots available, ten research on using only one AI chatbots available, and four research on integration AI chatbot available with social media application or platform or self-developed virtual environment. We will analyse how the different types of AI chatbots influence teachers' and students' perspectives and learning outcomes.

Large Language Model (LLM) or Generative AI (Gen AI) Chatbots

Firstly, we investigate 17 research studies on the learning outcomes of AI chatbot applications involving the Large Language Model (LLM) or Generative AI (Gen AI), such as ChatGPT 3.5, 4.0, and Bing Chat in language learning. This 17 research contains diverse studies, which are six research on both students' perspectives and

learning outcomes, three research about teachers' perspectives, one that involves both teacher's and students' perspectives in ChatGPT application in writing protocols, and one about using ChatGPT to create content for Master level in Law English by using ChatGPT.

Among the six research studies about students' perspectives on using ChatGPT, two have to do with learners' attitudes toward ChatGPT, and the other three research studies on students' acceptance and adoption of ChatGPT, where two research studies focus on informal digital language learning. Apart from that, there are six research regarding learning outcomes, where two research show that ChatGPT improves learners' writing skills; one is using ChatGPT in Chinese as a second language learning after school, and the other is using ChatGPT for formal writing. One research study illustrates that ChatGPT enhances only English vocabulary learning, and one research study underscores that ChatGPT improves English writing, grammar, and vocabulary acquisition. One research that integrates ChatGPT with AR glasses shows that it has the potential to enhance language learning. However, only one research study has shown that AI-generated feedback did not result in superior linguistic progress compared to feedback from a human tutor.

On the other hand, there is one research study that involved both teachers and students, which compared the experimental group (ChatGPT writing instruction protocol (CGWIP) with a control group (task-based model of L2 process writing instruction). The result illustrates that CGWIP significantly improved learners' writing skills and enhanced teachers' self-efficacy. Next, the three research involve teachers' perspectives, where one is about teachers' perception of ChatGPT academic integrity, one about teachers' beliefs in integrating ChatGPT and one about factors influencing teachers' acceptance and adoption of ChatGPT in their English Language Teaching.

Self-developed AI Chatbot with and without the Integration of Social Media Application or Platform

There are 15 studies on self-developed AI chatbots integrating social media applications or platforms. Facebook Messenger, Telegram and Line are the most used social media applications or platforms, with four studies each. WeChat and Google Assistant integration each has one research. There are three studies on Facebook Messenger integration, where one chatbot was built using either ActiveChat.ai or Google's Dialogflow only. The other is developed with both ActiveChat.ai and Google's Dialogflow to build and support a storybot for digital storytelling. The remaining research integrated Facebook Messenger with a self-developed AI chatbot and iReview. One research study used Google's Dialogflow to develop an AI chatbot that integrates with Google Assistant, which operates on a Google Home Mini device.

Meanwhile, for four research studies on Telegram integration, one research study designed a self-developed AI chatbot as part of the V-TREL system. The V-TRELis built on top of our proposed architecture to offer vocabulary training exercises generated from the common-sense knowledge-based ConceptNet and– in the background– to collect and evaluate the learners' answers for extending the ConceptNet with new RelatedTo relations for the trained words (Rodosthenous et al.,2020). The other three were researched with Telegram integration; one built a Reinforcement Learning Chatbot for the Kazakh language, one developed an AI chatbot for prosodic cues in Italian, and one developed an AI chatbot for sentence correction, vocabulary guessing and discussion purposes.

There are four studies on integration with LINE, three from a research project in building an AI chatbot called GengaBot for Japanese beginner Indonesian learners. One is developed via Python programming, based on rote learning for memorising vocabulary and a personal response system (PRS) to help with memory retention.

Among 26 studies on the implications of self-developed AI chatbots without integrating social media applications or platforms, ten research apply Google's Dialogflow machine learning AI platform, two using BlenderBot 3 and one using the Danbee AI platform. These are the open-source chatbot builder that has Machine Learning (ML) capabilities, which are available. Meanwhile, five studies have developed AI chatbots with virtual environments. The other six research AI chatbots are developed with different types of platforms, such as fuzzy neural network (FNN) (Gan, 2024), Figma (Setiawan et al., 2022), the combination of web applications to build AI chatbot (Ruan et al., 2021; Polyzi & Moussiades, 2023; Wu et al., 2023), string-matching method (Kharis et al., 2022). The remaining five research are self-developed from scratch, which are Argumate (Guo et al., 2023), Odinga teachable agent (Lee et al., 2023), SignGuru AI chatbot (Paudyal et al., 2020), Bookbuddy (Ruan et al., 2019) and CSIEC (Computer Simulation in Educational Communication) system (Jia, 2009).

Most of the research findings found that integrating social media applications or platforms in AI chatbots enhances students' language performance, learning engagement, and affective factors, such as learning motivation, learning anxiety, interest, and willingness to communicate. Language learning performance includes language skills, such as speaking, writing, reading and listening, as well as vocabulary and grammar proficiency. Nine research revealed that AI chatbot can improve students' speaking proficiency, eight research on enhancing vocabulary learning, five about improving grammar proficiency, two for writing skills and one for reading skills. Besides that, 14 AI chatbots showed that learners' sense of engagement was increased compared to their traditional class when using AI chatbots. In addition, nine research studies show that AI chatbots make learning fun, motivate learners to learn, improve their confidence and willingness to communicate, and decrease their anxiety.

Application of One and More than One AI Chatbot Application

Twenty-three studies use at least one and more than one AI chatbot application, whereas nine studies use only one AI chatbot application, and 14 studies use multiple AI chatbots.

Among the nine studies that apply only one AI chatbot application in language teaching and learning, the most used AI chatbots are Replika, Andy, and Cleverbot, with two studies each. Meanwhile, Mike, SHRDLURN and Smart Sender Platform each have one research. Seven research studies revealed that the available AI chatbots in the market have implications for students' learning outcomes. Three studies were found to enhance learning performance and speaking skills, and one research study was discovered on writing skills. For affective factors, two research studies are about enhancing learning motivation and related to learning anxiety. There is one research related to willingness to communicate (WTC). There are two research on factors affecting students' application of AI chatbots in language learning. On the other hand, among the 14 studies that apply multiple chatbots in language teaching and learning, we found that most AI chatbots researchers like to use are Andy, Mondly, Replika, Mike, Duolingo, Google Assistant and Memrise. Nine research has to do with language proficiency and affective

factors. Four research studies are about factors influencing students' learning with AI chatbots. Based on the nine studies on applying only one AI chatbot application and 14 studies on applying more than one AI chatbot, it was discovered that most AI chatbot studies are focused on English language teaching and learning. There is an exception, as there was only one study on students' perspectives when AI chatbots were combined with a 3D educational environment in Spanish as a foreign language. Of the nine research studies, only one does not specify which language is taught in the AI chatbot, while two of the 12 do not mention the language taught in the AI chatbots.

Application of AI Chatbots with Integration into Social Media Applications or Platforms and Others

There are two research studies on applying AI chatbots with Integration in Social Media Applications or Platforms in teaching and learning English, one from teachers' perspective and the other on high school students' learning outcomes. One research study on AI chatbots does not specify the types of applications used in language teaching and learning. The research mainly focused on the correlations between students' understanding of AI chatbots, their engagement in using these technologies, their level of technological competence, and their achievements in language learning (see Table 7 and Appendix).

Table 7. Implication of Different Types of Chatbots in Language Learning

No.	Authors	Implications
1.	Cai et al., 2023	Factors influencing learner attitude in ChatGPT-assisted language learning: -Information system quality and hedonic motivation have greater predictive power over perceived satisfaction and expectancy than self-regulation. -BI predicts learning effectiveness better than perceived satisfaction and performance expectancy.
2.	Cong-Lem et al., 2024	Teacher's perspectives on ChatGPT academic integrity: -Academic dishonesty in using ChatGPT: plagiarism, lack of originality of ideas, and use of AI-generated text without proper attribution. -Key factors driving students to cheat with AI tools potentially are perceived to be poor motivation and learning attitudes, combined with intense pressures to achieve academically. -Over-reliance on AI to generate content and ideas seriously hinders the development of students' genuine skills, critical thinking, and deeper language competencies. Hence, suggest for improvement: -advocate for increased regulations, the implementation of AI-based plagiarism detectors, and education on responsible AI use. -It is crucial to adapt language teaching pedagogies and assessments to incorporate personalised learning and process-oriented teaching approaches that support critical thinking and genuine learning motivation.
3.	Escalante et al., 2023	-AI-generated feedback did not result in superior linguistic progress among ENL students compared to those who received feedback from a human tutor. -Half the students preferred receiving feedback from a human tutor, and half preferred AI-generated feedback.
4.	Gao et al., 2024	EFL university teachers' beliefs in integrating ChatGPT: -Several concerns: neglecting traditional learning resources, academic integrity, and excessive reliance. -Influenced by previous experiences with LLMs, frequency of use, and self-evaluation on stages of LLMs integration. - the availability of IT personnel and their evaluation of IT infrastructure has no significant correlation to the university teachers' beliefs in integrating ChatGPT.
5.	Karataş et al.,	-ChatGPT positively affects students' learning experiences, especially in writing, grammar, and vocabulary

No.	Authors	Implications
	2024	acquisition, and enhances motivation and engagement through its versatile and accessible nature in various learning activities.
6.	Kostikova et al., 2024	- a new course for Law English training, a curriculum, a syllabus and a textbook using content created by ChatGPT were readied.
7.	Lee et al., 2023	-The effectiveness and user acceptance of VisionARy highlight its potential to cater to the needs of second-language learners and enhance their language skills through a more engaging and personalized approach.
8.	Ghafouri et al., 2024	-ChatGPT-based writing instruction protocol (CGWIP) significantly enhanced teachers' self-efficacy compared to the task-based L2 process writing instruction model. -CGWIP significantly improved learners' writing skills, and these effects persisted over time.
9.	Li et al., 2023	-Each participant made a noticeable improvement in their Chinese writing scores when using ChatGPT after school and during the reversal phases. -ChatGPT played a crucial role in correcting errors and facilitating the development of complete sentence structures. -the students expressed a sense of empowerment through their interactions with ChatGPT.
10.	Liu et al., 2024	Factors for acceptance and adoption of GPT chatbots among Chinese EFL learners in informal digital learning of English (IDLE): -How does the PU of chatbots for IDLE emerge from hands-on experimentation with these tools? -How does BI use increase as learners negotiate chatbot affordances and constraints? -How does AU of chatbots for IDLE involve using these tools as tutors or conversation partners?
11.	Punar Özçelik & Yangın Ekşi, 2024	-students found ChatGPT beneficial for acquiring formal register knowledge but perceived it as unnecessary for informal writing. - the participants questioned the effectiveness of ChatGPT in teaching neutral registers.
12.	Qu & Wu, 2024	Learners' perspectives on the adoption of ChatGPT - ChatGPT significantly amplified students' intrinsic motivation during their ESL learning. -Boredom, joy, focused immersion, and control emerged as significant mediating factors for the link between PEOU and BI.
13.	Shaikh et al., 2023	- ChatGPT is an effective tool for formal English language learning.
14.	Xiao & Zhi, 2023	-ChatGPT is a learning partner or personal tutor that provides personalized, easily accessible, and adaptive feedback. -ChatGPT assists in improving language proficiency when students apply critical thinking skills, such as modifying prompts, training the model, and verifying and selectively accepting its outputs. -ChatGPT facilitates idea generation for brainstorming purposes. -students demonstrated the capacity to think critically about the information generated by ChatGPT and reported their ability to modify prompts, train ChatGPT, verify, and selectively accept the information provided.
15.	Zhang & Huang, 2024	-employing an AI Chatbot based on LLMs aids students in acquiring both receptive and productive vocabulary knowledge in their English language learning. - Chatbots contribute to the long-term retention of productive vocabulary and facilitate incidental vocabulary learning.
16.	Dehghani & Mashhadi, 2024	Factors influencing acceptance of ChatGPT among EFL teachers in English Language Teaching -Significant relationship found between PEOU and PU; PEOU and BI to use; PU and BI to use; PSQ and PU; PSQ and PEOU; OCD and PU; PSQ and PU; PE and PEOU; PE and PU; PE and BI to use; PSE and PU; PSE and PEOU; SN and PU; SN and PEOU. -OCD and PEOU have no significant correlation.
17.	Liu & Ma, 2024	Using ChatGPT in Informal Digital Learning of English -PEOU fails to predict learners' attitudes, but PU is the full mediator for Attitude. -Learners' positive attitudes toward ChatGPT's usefulness tend to demonstrate a higher level of BI to use, which positively and significantly predicts their AU.
18.	Lee, et al,	- improves the speaking ability of the agent as the learner's skill improves.

No.	Authors	Implications
	2023	-Implement a character chatbot to increase students' learning efficiency by escaping the boredom of repeating learning and making it more like a human conversational environment. -apply to motivate students or preschoolers who are not interested in learning through role switching to participate voluntarily in learning. -improve the immersion of the educational effect.
19.	Berns et al., 2018	The benefits of VR chatbot: -Learners get the opportunity to experience the target language in real-world situations, and their actions of receiving real-time feedback facilitate them in revising and correcting their actions and answers.
20.	Cai et al., 2020	-Digital theatre with "the mechanisms of real-time spoken language evaluation and interactive switching of scenario & virtual costumes" has a better effect on language learning as students are more willing to learn the learning materials and more seriously participate in the dramatic learning processes. -Students are more enthusiastic when performing in drama as they can wear digital theatre costumes and are more willing to play the roles and read the scripts. -Students who experience the chatbot in digital theatre will try their best to speak the scripts loudly, clearly, and standardly to make the chatbot recognize their pronunciation and grammar successfully and achieve a sense of achievement.
21.	Divekar et al., 2022	-Students display statistical significance and retain improvement in Chinese as a Foreign Language (CFL) vocabulary, comprehension, and conversation skills. -Students were fun and engaging and found to create a generally more comfortable and less anxiety-ridden space for them to practice using the target language. -Benefits for the FL students: improvements in students' willingness to communicate in the target language, culture absorption, incidental learning, learning out of need, and other benefits of actual in-country immersion programs without the hassle.
22.	Gan, 2024	-FNN provides learners with organized chances to enhance language skills essential for effective interactions in the tourist sector while balancing correctness and fluency. -spoken English teaching is more effective when using an FNN to evaluate the efficiency of teaching English.
23.	Guo et al., 2024	- Students and Argumate formed a close partnership in a writing community where the chatbot assisted the learners in scaffolding their writing process. - students' writing process and use of technology to improve writing quality were connected to meeting task requirements and proficient essay conventions.
24.	Jia, 2009	- Users' behaviour when using CSIEC indicates that they prefer free chatting without spelling and grammar checking. - The increased percentage of long and long chats shows that the quality of free chats at CSIEC is improving.
25.	Kharis et al., 2022	-Gramabot assists students with a better understanding of the grammar taught in the course. - no clear sign that more intense prior knowledge of chatbots might connect to a better assessment of Gramabot.
26.	Kim & Su, 2024	-The experimental group demonstrated significant improvements in WTC, reduced anxiety levels, and enhanced communication confidence among students in quantitative analysis. -students' interviews revealed their experience of decreased anxiety, enhancing their WTC in Korean.
27.	Kim et al., 2022	-Ellie has the potential to become a practical language-learning companion in providing valuable opportunities to learn and practice an L2. -Students' ability to maintain a relatively lengthy conversation in English and engage in L2 problem-solving tasks with a substantial amount of meaningful negotiation when they communicate with Ellie.
28.	Kohnke, 2023	- Chatbots are helpful in L2 instruction, as they can motivate students to complete out-of-class work, prepare for class more confidently, and permit them to study whenever and wherever they wish. - using a chatbot in the classroom may distract students from the task at hand, increasing the teacher's classroom management responsibilities.
29.	Kwon et al., 2023	-AI chatbot-based writing practice had a facilitating effect on students' writing test performance. -AI chatbots help improve students' language skills and make them feel comfortable when learning a foreign language.

No.	Authors	Implications
30.	Lee et al., 2024	-Experimental group, each with a presentation task and interaction with AI chatbots in 3D metaverse environments had positive effects on all facets of Global Englishes (GE), e.g., acceptance of one's local English, acceptance of other Englishes, native-speakerism, ELF confidence and intention, and willingness to incorporate GELT into teaching). -The AI chatbot task exerted a stronger effect on English as a lingua franca (ELF), confidence, and intention.
31.	Liang et al., 2023	-in chatbot-driven language-learning tools, corrective feedback is more effectively delivered through means other than the social chatbot, such as a GUI interface. -guided self-correction offers a superior learning experience compared to providing explicit corrections, particularly for learners with high learning motivation or lower linguistic ability.
32.	Paudyal et al., 2020	- The execution performance of 26 new learners for 14 different American Sign Language in AI chatbots increased from an average of 40% to 68%.
33.	Polyzi, & Moussiades, 2023	-AI chatbot improved students' vocabulary learning performance. -Students considered AI chatbots over traditional learning methods as they are enjoyable, easy to use, and help them learn vocabulary.
34.	Qian et al., 2023	- A dialogue system with curriculum infusion improves students' understanding of specific English words and phrases and increases their interest in practising English.
35.	Ruan et al., 2021	-Student's fluency improved more with EnglishBot. - EnglishBot users showed higher engagement and voluntarily spent 2.1 times more time interacting with EnglishBot. -Conversational interfaces may benefit foreign learners' oral language learning, particularly in casual learning settings.
36.	Ruan et al., 2019	-Children enjoyed speaking English with our virtual tutoring chatbot and were highly engaged during the interaction. -Children were engaged more than half of the time while using BookBuddy.
37.	Pham et al., 2018	-Users still interact positively with the AI chatbot, and even though it is difficult to use, it should be practised before use. -The AI chatbot is still designed with low depth, or it misses users' questions, so they end the conversation soon after the bot does not meet the requirements.
38.	Setiawan et al.,2022	-AI chatbot facilitates students in learning English, memorising vocabulary, understanding grammar, and increasing confidence in speaking English.
39.	Wang et al., 2017	-AI chatbot and time machine facilitated embodiment for learners who feel somewhat natural by engaging with these learning artefacts, which increases the learners' sense of immersion and presence. -Real-time interactions with the chatbot and the time machine provided a sense of authenticity, which determined to what extent the learner could transfer what he/she learned in the 3D virtual world to real life.
40.	Shin et al., 2024	-Customized AI chatbots could offer corrective feedback (CF) when students make non-target utterances and elicit learner uptake successfully.
41.	Wu et al., 2023	-Numerous features in SpokenBot straighten oral skills by expanding the vocabulary bank, increasing pronunciation accuracy, enhancing speaking fluency and improving presentation skills. - Portfolio management raises learning motivation and continuous self-learning. - Systematic learning is provided to support appropriate study planning, which raises learning motivation and continuous self-learning,
42.	Yang et al., 2022	-AI chatbots encourage students to engage in conversation, which rarely occurs in general EFL classes in Korea. - Chatbot offered opportunities for the students to practice the target language using achievable task goals.
43	Kohnke, 2022	-AI pedagogical chatbot provides support with human-like interactions, which enhances students' sense of engagement. - their interactions with the chatbot eased their sense of isolation during lockdown, which positively impacted their learning.
44.	Chiaráin &	-AI chatbot increase motivation and promotes positive attitudes towards the language.

No.	Authors	Implications
	Chasaide, 2016	-ABAIR text-to-speech (TTS) was given a high approval rating as it is proven 'fit-for-purpose' in which high-quality synthetic speech can play a significant role in language learning.
45.	Bailey & Almusharraf, 2021	- directives requesting user input resulted in 35% more output, indicating students took more action when told to do something than when asked. -Chatbot-to-user directives resulted in the most content produced. - For novelty effect, fewer students volunteered to do the sixth and final chatbot activity, but those who did produce a word count on par with their initial chatbot activity.
46.	Barley et al., 2021	- When measuring student levels of participation and their perception of storybot interaction, it was found that students who were confident using storybots sent many messages. -With interest, students' self-rated L2 proficiency correlated positively with participation rates. -students read nine times more than they wrote, indicating a high degree of reading comprehension necessary for storybot interaction. -Storybot benefits L2 comprehension, as measured by cohesion between dialogue nodes or even built-in chatbot quizzes.
47.	Chen et al., 2020	-The AI chatbot significantly improved students' learning achievement, and having a one-on-one environment may lead to better outcomes than what could be achieved in a classroom. -PU was the predictor of BI, whereas PEOU was not. -The AI chatbot can enhance students' learning of Chinese vocabulary.
48.	De Iacovo et al., 2021	- a chatbot as a proactive learning support to improve oral skills in Italian L2.
49.	Haristiani et. al., 2022	-helpful in enhancing students' basic Japanese grammar learning, improving vocabulary mastery related to vocational terminologies, and providing practices for basic Japanese language level exercises. -attractive and interactive features foster learners' autonomy and independence learning due to its practicality, portability, accessibility, and flexibility.
50.	Haristiani & Rifa'I, 2020	-Gengobot is practical, user-friendly, innovative, and helpful in supporting students' Japanese language learning, specifically Japanese grammar learning. -Gengobot integrated with LINE social media is highly accessible, indicated to improve students' motivation in learning Japanese grammar, and adequate to enhance students' environment.
51.	Haristiani & Rifa'I, 2021	-Gengobot is an exciting and innovative medium to support Japanese autonomous learning so learners can decide how they learn using this application to improve their Japanese grammar skills. - Gengobot is more interactive than other Mobile-based media, making learners more interested in using Gengabot as a Japanese grammar learning medium.
52.	Johnson et al., 2022	-Students are satisfied when using Escapeling to learn vocabulary and grammar. -Escapeling has the potential as a collaborative learning environment.
53.	Khlaisang & Sukavatee, 2023	-are flexible, allow for learners' self-paced learning activities, and encourage interactivity and knowledge sharing among users. -PU influenced perception towards AI chatbot while facilitating conditions influenced PEOU. -Organised learning using an AI chatbot application develops students' listening and speaking skills, which they have more significant skills when communicating in English. -learning with an AI chatbot can support cooperative, collaborative, active, constructive, creative, and social learning, increasing students' motivation to learn well beyond the limitations of learning in standard classrooms.
54.	Kohnke, 2023	-students enjoyed interacting with the chatbot both in and out of class and perceived that it improved their English skills. -Chatbot helped students observe and correct language-related errors. -the feedback they received from the chatbot provided comprehensible input and an opportunity to modify language output.
55.	Rodosthenous et al., 2020	-the results show a tiny yet positive shift in learners' vocabulary skills after using LingoGameBot.
56.	Oralbayeva et	-The RL chatbot did not result in significant learning gains compared to the non-RL chatbot.

No.	Authors	Implications
	al., 2022	-non-RL chatbot gained equally positive feedback as the RL chatbot. -The RL chatbot had a better trend because participants gave overall more positive responses than its non-RL version, as a feature in the RL chatbot was perceived as more significant and associated with its knowledgeableness.
57.	Wu et al., 2024	- learners engage in more rounds of dialogue and use more interaction strategies when interacting with IPAs than with their peers. - The latent Dirichlet allocation (LDA) analysis showed that the communication topics between the learners and the IPA are rich. Conversations with IPAs can significantly improve learners' listening and speaking abilities.
58.	Lee & Jeon, 2022	-the majority of students identified human elements in the voice-controlled conversational Agent (VCA) perceives it as a human-like partner or something between an artefact and a human. A strong tendency toward anthropomorphism indicates VCA's great potential as an interactive language partner in EFL contexts.
59.	Alrajhi, 2024	-AI chatbot is perceived as supportive of L2 practice and writing development, is interesting, enhances motivation, and alleviates writing anxiety. -demotivating factors: chatbot's interactional and instructional abilities, including the lack of extended conversations, sensitivity to inaccurate language forms, and sporadic irrelevant responses. -L2 proficiency does not affect overall views on the chatbot-mediated interaction, except for the usefulness aspect for L2 practice, which has significantly more positive views from high-intermediate students.
60.	Çakmak, 2022	-Student performance with Replika was significantly better than their face-to-face peer interactions. -there were negative perceptions and attitudes toward the chatbot interaction as students reported facing difficulties in being understood precisely, which might have contributed to higher anxiety in L2 speaking. -Overall, although chatbot interaction is a novel way to provide speaking practice for students, interaction with a chatbot might not be a reliable way to lessen their anxiety with L2 speaking.
61.	Duong & Suppasetseree, 2024	-Students showed significant improvement in English speaking skills after using an AI voice chatbot to practice speaking. - English speaking skills improved after the intervention because they could speak English better using suitable hedging words, grammar structures, and vocabulary.
62.	Fathi et al., 2024	- AI-mediated interactive speaking activities were more effective in improving EFL learners' speaking skills and WTC. -Students had positive attitudes and perceptions towards the AI-mediated speaking instruction.
63.	Fryer et al., 2017	- Comparisons of task interest under different partner conditions over time indicated a significant drop in students' task interest with the Chatbot but not the Human partner. - only task interest with the Human partner contributed to developing course interest.
64.	Fryer et al., 2019	-prior interest in human conversation partners was the best predictor of future interest in chatbot conversations. -Prior language competency was more strongly linked to interest in chatbot than human conversations. -the qualitative experience of having "learned more" with the chatbot was strongly connected to task interest, even when reporting communication difficulties.
65.	Lin & Mubarok, 2021	-The mind map-guided AI chatbot approach (MM-AI) promoted the students' English-speaking performances more than the conventional AI chatbot approach (C-AI). -MM-AI promoted the students' learning performance and organized the interaction between the robots and humans more than the C-AI did.
66.	Mu & Sarkar, 2019	- when playing with an AI chatbot, students tend to achieve the same or better task performance without detriment to user experience: in fact, participants reported less effort and higher performance in the restricted condition. -students who organically developed simpler languages tended to perform better. -students in restricted conditions are forced to be perfectly consistent, which improves model learning. - a guided, consistent language helps students understand the system's limitations and, within these constraints, infer the abstractions needed to succeed in the task.

No.	Authors	Implications
67.	Nozhovnik et al., 2023	-AI chatbot increased student motivation, improved their English language proficiency in reading and language usage, and enhanced engagement. -Students appreciated this chatbot-based tool's ease of use and usefulness in improving their English language skills.
68.	Annamalai, et. al., 2023	-chatbots can improve English language learners' competence, autonomy, and relatedness -Chatbots lack an emotional environment and give inaccurate information about learning English. -emphasises the significance of providing language learners opportunities to interact with and practise the language in a supportive and non-judgmental environment. - Learners who use chatbots gain confidence and proficiency in using the English language and have a stronger connection to the language and the learning process.
69.	Annamalai, et. al., 2023	-Performance and Effort Expectancies contribute to the positive experience of using Chatbots for language learning, hence serving as the Pull factors. -Social Isolation contributed by the sense of robotic interaction, emotionlessness, and lack of flow in conversation serve as the Push Factor in using Chatbots for language learning.
70.	Annamalai, et. al., 2023	-Three of these- passive, active, and constructive modes- gave students exposure to language materials via meaningful AI chatbots activities, contributing to their overall proficiency in the subject matter. -interactive learning through AI chatbots had issues with responsiveness and lacked personalized feedback, hindering language skill practice and critical thinking.
71.	Belda-Medina & Calvo-Ferrer, 2022	-positive results regarding perceptions concerning the integration of conversational agents in language learning, particularly concerning PEOU and attitudes (AT), but the scores for BI were more moderate. -quantitative analysis demonstrated that gender and educational setting did not affect participants' satisfaction with the linguistic level of the conversational partners; qualitative findings unveiled some gender-based differences regarding customizing options and topics of interaction. -Female participants were more attentive to using inclusive design and language and more assertive about gender stereotyping. - Multimedia content, gamification and use of non-verbal language in AI chatbots could become critical factors for learners' satisfaction, as indicated by the teacher candidates
72.	Belda-Medina & Kokošková, 2023	-The Language Experience dimension (LEX), including elements such as Semantic Coherent Behaviour, Sentence Length and Complexity, and Speech Recognition and Synthesis, revealed that none of the AI chatbots reached a moderate level of satisfaction among EFL teacher candidates. -The Design Experience dimension (DEX) underscored the importance of user-friendly interfaces and engaging multimedia content in fostering user engagement and satisfaction. - Enhancements in adaptive user interfaces and the incorporation of social media and emerging technologies are needed to simulate human-student interaction and enrich the language learning experience.
73.	Coniam, 2008	-The current generation of AI chatbots cope best when presented with one-clause questions or statements embracing straightforward propositions with minimal cohesive linkage to previous utterances. -Some AI chatbots have large lexicons and can mention many facts, which indicates extensive world knowledge. All the chatbots can do is mention specific facts and issues. -AI chatbots' 'knowledge' is, to an extent, a sham since they cannot incorporate this knowledge in follow-up discussions.
74.	El Shazly, 2021	-students experienced certain anxiety levels at similar percentages pre- and post-intervention, which compels the learners to worry about failing the course rather than promoting their foreign language learning. -students' anxiety levels were not alleviated when employing non-threatening AI applications; in fact, Foreign Language Anxiety (FLA) increased slightly.
78.	Fryer & Carpenter, 2006	- Most students enjoyed using the AI chatbots. -Students generally felt more comfortable conversing with the bots than with a student partner or teacher.
79.	Kemelbekova, et. al., 2024	-Both experimental and control groups showed a significant improvement in their oral communicative abilities, as evidenced by their performance in the two speaking tasks, namely reading a text aloud and answering

No.	Authors	Implications
		questions. -There is no significant difference between the two groups regarding pronunciation, but a discernible variation was detected in their intonation and stress patterns during the reading task.
80.	Kim et al., 2021	- Both experimental and control groups indicated that both groups showed significant improvement in the two speaking tasks, including reading a text aloud and responding to questions. -The students improved their pronunciation, intonation, and stress both in the low- and intermediate-level groups. -There was no significant difference in the low-level student group for fluency measured in WPS in the read-a-text-aloud task with AI chatbots. At the same time, there was a statistical difference in the intermediate-level student group. -There was no significant difference in pronunciation between low-level and intermediate-level pronunciation in the read-a-text-aloud task in AI chatbots. However, the two group levels show a notable difference in intonation and stress.
81.	Mohamed & Alian, 2023	- Students can utilize AI chatbots without the help of the instructors, which in turn encourages them to become autonomous learners. - students believed that the AI chatbots might simulate an interaction cycle so they could practice the target language. -students felt that the AI chatbots boosted their enthusiasm and confidence, which ultimately helped them feel active and more comfortable.
82.	Alm & Nkomo, 2022	-Results indicate a degree of curiosity and a willingness to engage in a conversation with chatbots. -students expressed frustration if the dialogues did not correspond to their learning goals or if they were excluded from using the bots because of technical or payment issues or discontinuation of service.
83.	Chuah & Kabilan, 2021	-teachers perceived the use of chatbots in giving feedback to their students as very helpful, though some needed extra training on how to use them. -teachers thought chatbots could simulate an interaction cycle for students to practice the target language. -Teachers believed chatbots augmented a more significant social presence, eventually creating an environment where their students could be active.
84.	Ye et al., 2022	-grammar and pronunciation accuracy in students' oral English improve after language practice with Xiaoying -most learners have a positive attitude towards this chatbot as a language learning tool to improve oral accuracy.
85.	Lorenzo et al., 2013	-Results show that the possibility of cooperating and collaborating in an explicit social context, such as a 3D educational environment, in combination with enhanced communication tools (chat, video chat or VoIP) and intelligent assistants (chatbots or NPCs) play a pivotal role in user acceptance of MMOL platforms. -the most crucial determination of MMOL platform adoption seems to be the perceived value of cooperation, collaboration, communication and in-world assistance on the MMOL platform. - The high regression coefficient value associated with the Communication and Collaboration Capabilities (CCC) variable means that this constructor is a significant antecedent to perceived usefulness. -AI chatbot performance directly affects instructional designers and educational institutions because the analysis suggests that communications and assistant capabilities, playfulness and community ideas are crucial for developing, designing and assessing MMOL platforms.
86.	Tai & Chen, 2020	-Google Assistant significantly promoted EFL students' WTC, enhanced communicative confidence, and reduced speaking anxiety. -students enjoyed playing games with Google Assistant and talking to chatbots, which helped them feel less anxious and motivated to use English for real and meaningful communication. -IPA-based interaction provided a less threatening environment in which students displayed higher levels of engagement, motivation, confidence, and WTC in the target language.
87.	Zhang et al., 2024	-There is a significant correlation between students' understanding of AI chatbots and their language learning achievements, with practical involvement mediating this relationship. - technological proficiency was a significant moderator in influencing the link between knowledge and results.

Discussion

The State-of-art Topics related to Using AI Chatbots in Language Teaching and Learning from 2006 to 2024

The overall state-of-the-art review of AI chatbot applications for language education shows a vast increase from 2020 onwards in publication and citation in WOS and Scopus. Hence, the increment in publications and citations in 2020 can be caused by the COVID-19 pandemic lockdown, where almost all physical economics activities could not be carried out, as did the physical classes at any education institute. This restriction increases the demand for online learning, which requires the development of new technology applications and devices so that language teaching and learning can still be carried on. Hence, the demand for AI chatbots had increased during this period. Although the pandemic was ending in 2022, Open AI launched its latest generative AI, ChatGPT, which caused the AI chatbots to continue to develop.

Looking in detail at the trends of AI chatbots between 2006 and 2023, we can see that the amount of research before 2020 was limited. From 2006 until 2015, there is one self-developed AI chatbot without any social media application or platform integration and three research on using multiple AI chatbot applications. Jia (2009) published the first self-developed AI chatbot for EFL learners, especially Chinese speakers who want to learn English. According to Fryer and Carpenter (2006), this scenario happened because most of the AI chatbots available in the marketplace are for native English speakers. Since the AI chatbots available in the marketplace have unique features, which provide advantages and disadvantages in language learning design, most research combines more than one AI chatbot during their language teaching. One research study combines AI chatbots with a virtual world called OpenSims, which is available to teach Spanish as a foreign language.

From 2016 until 2019, there has been an increase in self-developed AI chatbots without integration of social media applications or platforms, and the application of only one AI chatbot in language teaching has started in teaching English as a foreign language. Meanwhile, the number of studies delved into the virtual world has increased for self-developed AI chatbots. This application has also started to apply in other foreign languages, such as German (Bern et al., 2018) and the endangered Irish language (Chiaráin & Chasaide, 2016). However, according to Pham et al. (2018), when they applied their self-developed AI chatbot to Vietnamese who are learning English, they found that their current development was still designed with low depth or that it did not catch the questions of students that they ended the conversation soon after the bot did not meet the requirements. Although students perceived the AI chatbot as challenging, they still interacted positively. Hence, they suggested that students should be practised using an AI chatbot to familiarise themselves before using it for language learning.

Between 2020 and 2022, the authors see that all AI chatbots either self-developed with or without integration in social media applications or platforms, nor the use of one or more than one AI chatbots available in the marketplace with and without social media integration has increased. This phenomenon happened due to the impact of the COVID-19 lockdown, where physical education was restricted. All the teaching and learning were shifted to online learning that required devices and specific applications and platforms, which led to the maturity of the development of AI chatbots. At the same time, a wide variety of research focuses on teachers' and students'

perspectives, learning outcomes, and learning affection. Since some of the AI chatbots available in the marketplace did not fulfil the needs, especially in teaching English to those whose native language is not English, there are six self-developed without social media integration and five self-developed with social media integration. During that time, there were self-developed AI chatbots for learning foreign languages, such as Mandarin, German, Japanese, Korean, and Italian, where the medium of explanation for that target language was the students' native language, which was not English. There is a need for self-developed AI chatbots as most foreign language learning AI chatbots are taught and explained in English, which is not the student's native language. Moreover, a self-developed AI chatbot called SignGuru allows students to learn American Sign Language words and phrases.

At the end of 2022, OpenAI introduced ChatGPT (Generative Pre-trained Transformer) and gained popularity as it is an artificial intelligence chatbot that demonstrates the ability of digital devices to perform the tasks inherent to intelligent beings. It can create content, generate texts, answer questions and perform other Natural Language Processing tasks when the user asks (Kostikova et al., 2024). Hence, there is an inclining number of research on the effectiveness and implications of ChatGPT from the perspective of teachers and students. Although ChatGPT attracts the eye of teachers and students, the number of self-developed AI chatbots that do not integrate social media applications or platforms continues to increase. The same situation also happens when applying only one or multiple AI chatbots found in the marketplace. However, the number of research on integrating self-developed AI chatbots with social media applications or platforms and the AI chatbots available with social media has declined. This situation may be due to the lifting of lockdown restrictions, where every physical activity, including physical classes, has returned to normal. There is one research on students' general perspective on using AI chatbots in language learning. Despite there is the trend of ChatGPT application in language teaching and learning, ChatGPT's design is to provide information and perform tasks according to the input prompt, so it is limited in its application to language education based on everyday English usage (Lee et al., 2020; Shin et al., 2024). Therefore, this can explain the increase in research in self-developed AI chatbots without integrating social media applications or platforms and the application of one or multiple AI chatbots available in the marketplace. The bibliometric analysis in Vosviewer and the systematic review share some similarities. LLMs or Gen AI chatbots, such as ChatGPT, English language teaching, language learning model, technology acceptance, second language writing, and Dialogflow, are the main trends of publications in 2023.

The Influence of Different AI Chatbots towards the Teachers' and Students' Perspectives on Language Teaching and Learning and Students' Learning Outcomes

Large Language Model (LLM) or Generative AI (Gen AI) Chatbots

Applying Large Language Model (LLM) or Generative AI (Gen AI) chatbots in language teaching and learning has positive and negative implications. Based on students' learning outcomes, research has shown positive effects on English vocabulary knowledge (Zhang & Huang, 2024), assisting learners in developing their English writing abilities, particularly in formal register (Punar Özçelik & Yangın Ekşi, 2024) as well as Chinese language writing skills by correcting erroneous characters and developing well-structured sentences (Li et al., 2023), improves language skills, particularly in writing, grammar, and vocabulary acquisition, and increases motivation and engagement for its versatility and accessibility in various learning activities (Karataş et al., 2024). However,

research by Escalante et al. (2023) showed no significant differences in students' English writing proficiency. Moreover, half of the students prefer Generative AI feedback, while half prefer human tutor feedback. The difference in results of this research can be explained by the study's design, where those studies get positive results, ChatGPT, or any LLM chatbot, which acts as a learning support tool. Moreover, students usually get guidance from teachers on how to use ChatGPT, and use them during class (Karataş et al., 2024; Punar Özçelik & Yangın Ekşi, 2024) as well as after school (Li et al., 2023). Meanwhile, the study by Escalante et al. (2023) applied quasi-experimental research, where the experimental group only got feedback from ChatGPT, and the control group got feedback from the human tutor. ChatGPT can be more enhancing language skills once it integrates with AR glasses (Lee et al., 2023).

Applying the Large Language Model (LLM) or Generative AI (Gen AI) chatbots in language learning still requires teachers' guidance or instruction. Hence, teachers play an essential role in applying Gen AI chatbots; their perspectives are critical. Cong-Lem et al. (2024) research shows that teachers reveal their concerns about AI-based academic dishonesty, particularly involving generative chatbots like ChatGPT. The core forms of academic dishonesty are plagiarism, lack of originality of ideas, and using AI-generated text without proper attribution. The key factors that cause students to cheat with AI tools are poor motivation and learning attitudes, combined with intense pressures to achieve academically. They also elucidate their concern about students' over-reliance on Gen AI to generate content and ideas, which can seriously hinder the development of students' genuine skills, critical thinking, and deeper language competencies. However, Gao et al. (2024) discovered that teachers who have used LLMs have higher positive and lower negative beliefs than those who have never used LLMs. Hence, teachers with a higher frequency of LLM usage are likelier to believe LLMs can enhance students' academic achievement, be an effective and valuable instructional tool, and improve students' understanding of critical concepts and ideas. Therefore, teachers' perspectives in accepting and adopting GPT technologies require their own usage experience of the tools. Hence, PU and PEOU are pivotal determinants (Dehghani & Mashhadi, 2024). Teachers will be concerned about using ChatGPT as they perceive it is easy to use and useful in enhancing their English Language teaching performance, increasing their teaching efficiency, and producing better outcomes. Besides that, EFL teachers' satisfaction with ChatGPT's curriculum—considering its interest, diversity, appropriate difficulty, and adaptability to various user levels— strongly influences their perception of its usefulness.

Cong-Lem et al. (2024) reveal teachers' concerns about students' academic integrity and over-reliance on LLMs in language learning. However, according to Xiao and Zhi (2023), students exhibited critical judgment in evaluating the quality of ideas and outputs generated by ChatGPT and the ability to modify prompts to maximise learning benefits. Therefore, from the perspectives of students, the role of Gen AI tools can be (1) students' learning partner or personal tutor that provides personalised, easily accessible, and adaptive feedback; (2) to assist students in improving language proficiency by enhancing their critical thinking when they are modifying prompts, training the model, and verifying and selectively accepting its outputs; (3) to facilitate students in idea generation for brainstorming purposes. As students perceive LLMs as an essential tool for their language learning, they need to accept and adopt these applications in their language learning. Hence, Liu et al. (2024) have revealed three criteria for students' acceptance and adoption of GPT technologies in their informal digital learning of English (IDLE): (1) how perceived usefulness of AI chatbots emerges from hands-on experimentation with these tools;

(2) how behavioural intention to use increases as learners negotiate AI chatbot affordances and constraints; (3) how actual use of AI chatbots for IDLE involves using these tools as tutors or conversation partners. From these points, we can see that perceived usefulness, intention to use and actual use of AI chatbots play a vital role in convincing students to use GPT technologies. Cai et al. (2023) also support that behavioural intention predicts learning effectiveness in ChatGPT-assisted language learning better than perceived satisfaction and performance expectancy. Besides that, based on Qu and Wu's (2024) findings, boredom, joy, focused immersion, and control emerged as significant mediating factors of the link between perceived ease of use and behavioural intention. Furthermore, according to Liu & Ma (2023), students who have positive attitudes toward the usefulness of ChatGPT tend to demonstrate a higher level of behavioural intention, which positively and strongly predicts their actual use of this latest AI-powered tool in English learning outside the classroom.

Self-developed AI Chatbots without Integration of Social Media Application/Platform

Both self-developed AI chatbots with and without social media application/platform integration provide advantages, especially in language learning performance, learning engagement and learning affective implications. Language learning performance, including sign language, showed a positive impact, as Paudyal et al. (2020) found that the execution performance of students in American Sign Language increased with the use of the SignGuru chatbot. The language performance also includes language skills, such as writing, speaking, listening, reading, vocabulary, and grammar.

Different designs of AI chatbots will affect speaking performance differently, especially for self-developed AI chatbots without integrating social media applications or platforms. For example, Lee et al. (2023) designed a teachable Odinga agent that implements a character chatbot that students need to teach. This study shows that students improve their speaking ability as they need to get into a role switching to teach a chatbot, which increases their learning efficiency by escaping their boredom of repeating learning found in other chatbots. Apart from that, Cai et al. (2020) apply the concept of digital theatre in designing the AI chatbot, where students must try their best to speak the scripts loudly, clearly and standardly to make the chatbot recognise their pronunciation and grammar successfully. Wu et al. (2023) designed a SpokenBot containing portfolio learning that provides systematic learning to support students' appropriate study planning. The numerous features of SpokenBot can strengthen students' oral skills by expanding the vocabulary bank, increasing pronunciation accuracy, enhancing speaking fluency, and improving presentation skills. Ruan et al. (2021) designed a conversational interface AI chatbot that benefited foreign learners' oral language learning, particularly in casual learning settings. Gan (2024) developed an AI chatbot using a fuzzy neural network (FNN). Using student data, he used the AI chatbot to teach spoken English to tourists. The work process involves evaluating each student's language proficiency, gathering and preprocessing data, training the network, assessing its performance, integrating it into a tailored language learning platform, and regularly providing feedback and correction to help students improve their spoken English abilities. The same results were found in research on integrating AI chatbots with social media applications or platforms. Wu et al. (2024) used Mandarin Second Language Intelligent Personal Assistant (MSLIPA), an intelligent personal assistant integrated with WeChat, in their study. They found that the application significantly improves students' listening and speaking abilities by facilitating rich communication topics, encouraging them to have

more rounds of dialogue, and using more interaction strategies than their peers. De Iacovo et al. (2021) self-developed an AI chatbot integrated with Telegram and found that the chatbot is a proactive learning support to improve students' L2 Italian oral skills. Khlaisang and Sukavatee (2023) built an AI chatbot incorporated into Facebook Messenger for reading and writing, whereas an iReview system for listening and speaking. The researchers revealed that the organised learning design in the AI chatbot improved students' listening and speaking skills, which enhanced their communication skills. Kohnke (2023) discovered that students enjoyed interacting with the AI chatbot developed with Google's Dialogflow, integrated into Facebook Messenger, both in and out of class and perceived that it improved their English skills.

Meanwhile, for writing performance, according to Guo et al. (2023), Argumate chatbot enhances students' argumentative writing by providing essential assistance to learners, where the writing process with the aid of the chatbots helps learners improve their writing quality. Besides that, Kwon et al. (2023) self-developed an AI chatbot with encoding expressions from an elementary school English textbook for English writing practice. They found that AI chatbot writing provided more hands-on writing practice to individual learners, which offered students more effective and focused writing activities, allowing them to improve their writing. However, there is no research on the application of self-developed AI chatbots with integrated social media applications or platforms for writing skills. Moreover, for reading proficiency, according to Barley et al. (2021), AI storybots make students read nine times more than they wrote, indicating a high degree of reading comprehension necessary for storybot interaction.

Besides, vocabulary learning has positive implications when using self-developed AI chatbots with and without social media applications or platform integration. Without integrating social media applications or platform AI chatbots, Polyzi & Moussiades (2023) created a digital vocabulary assistant, which includes drag and drop, hangman and quiz, and chatbot as an English language learning assistant. The AI chatbot contains text-to-speech and speech-to-text assistive technologies that were used to facilitate the educational process and simulate communication to a higher degree to help better assimilate sociolinguistic skills. The study found that using the application for learning vocabulary statistically significantly outperformed student performance compared to traditional methods studies. Besides that, Qian et al. (2023) apply a user-adaptive generative chatbot for language learning. The result showed that the curriculum-incorporated chatbot help students learn specific words and phrases. Setiawan et al. (2022) developed an AI chatbot to help students facilitate English vocabulary memorisation.

Moreover, self-developed AI chatbots that integrate with social media applications or platforms also have positive impacts. For example, Chen et al. (2020)'s AI chatbot named Xiaowen, which was built via Python programming and integrated with LINE, was developed based on rote learning for memorising vocabulary and a personal response system (PRS) to help with memory retention. The study discovered that the AI chatbot improved students' learning achievement, especially in Chinese vocabulary learning. Moreover, the one-on-one environment in the AI chatbot may lead to better outcomes than what could be achieved in a classroom. Haristiani et al. (2022) revealed that with the integration of LINE in their self-developed Gengabot chatbot, students' vocabulary mastery related to vocational terminologies had been enhanced. There are two research that integrated

their self-developed AI chatbot with Telegram. One AI chatbot named LingoGameBot is a part of the V-TREL system that offers vocabulary training exercises generated from the common-sense knowledge-based ConceptNet, showed a minimal positive shift in students' vocabulary skills after using it (Rodosthenous et al., 2020). The other AI chatbot, Escapelling, is a sci-fi 'escape room' where students are satisfied using it to learn vocabulary.

Last, for language grammar, Shin et al. (2024) and Liang et al. (2023) both make their own AI chatbots with corrective feedback (CF) for English learning. Shin et al. (2024) installed a CF function using 'action and parameters' and 'define prompts' options in Google Dialogflow, a chatbot-building platform where students must respond to the chatbot in text and voice. The result showed that chatbots can provide CF properly when upper-grade EFL learners produce non-target utterances. Furthermore, students can make the correct reformulation of their last utterance that contained a grammatical error. Liang et al. (2023) created a synchronous CF and self-correction capabilities for grammar called Chatback using Blenderbot 3 3B. The result of the research discovered that the chatbot-driven language-learning tool's corrective feedback is more effectively delivered through means other than the social chatbot, such as a GUI interface. Furthermore, guided self-correction offers a superior learning experience compared to providing explicit corrections, particularly for learners with high learning motivation or lower linguistic ability. Although particular chatbots are not developed for corrective feedback, Kohnke (2023) said the AI chatbot helped students observe and correct language-related errors. Moreover, the feedback they received from the chatbot provided not only comprehensible input but also an opportunity to modify language output.

Furthermore, Kharis et al. (2022) built an AI chatbot named Gramabot by using the string-matching method. The chatbot functions as a medium for German grammar learning. The result revealed that Gramabot assists students with a better understanding of the grammar taught in the course. Moreover, Gengobot's design integrates LINE and is user-friendly and innovative. It supports Japanese autonomous learning so learners can decide how they learn using this application, which improves their Japanese grammar skills and is effective for learners to learn (Haristiani & Rifa'I, 2020; 2021; Haristiani et al., 2022). Escapelling chatbots satisfy students in learning grammar (Johnson et al.,2022).

On the other hand, self-developed AI chatbots can provide human-like interaction for learning engagement, enhancing students' sense of engagement (Kohnke, 2022). Some of the AI chatbots discovered that when using AI chatbots, there is an increment in students' percentage of having long and longer chatting time or more rounds of dialogue compared to a real human tutor, such as with CSIEC (Computer Simulation in Educational Communication) system (Jia, 2009), Chatbot Ellie (Kim et al., 2022; Yang et al., 2022), AI chatbot EnglishBot (Ruan et al., 2021) and an intelligent personal assistant integrated with WeChat called MSLIPA (Wu et al. 2024). In addition, the AI chatbot enriches the communicative topics that encourage students to use more interaction strategies than when interacting with their peers (Wu et al., 2024). Moreover, it also found that students who engaged with Bookbuddy, a scalable virtual reading companion that turns any reading material into an interactive conversation-based English lesson, engaged students more than half of the time in reading (Ruan et al., 2019). Furthermore, Bailey and Almusharraf (2021) also revealed that directives requesting user input in the AI chatbot resulted in 35% more output, indicating that students took more action when told to do something than when

asked. Hence, Chatbot-to-user directives resulted in the most content produced. The design of the conversational interface also plays a crucial role in enhancing the sense of engagement benefits foreign learners' oral language learning, particularly in casual learning settings (Ruan et al., 2021).

Besides that, some self-developed AI chatbots integrate with virtual environments, such as virtual reality (VR) (Wang et al., 2017; Berns et al., 2018), C# and Kinect's motion detection function for virtual costume module in digital theatre (Cai et al., 2020), and extended reality (XR) (Divekar et al., 2022), as well as voice-controlled conversational agent (VCA) (Lee & Jeon, 2022). The integration of virtual environments provides learners with the opportunity to experience the target language in real-world situations. Their actions of receiving real-time feedback facilitate them in revising and correcting their actions and answers (Berns et al., 2018). According to Wang et al. (2018) that based on learners' experience in VR chatbots, their sense of immersion was increased as they felt somewhat real through making engagement with learning artefacts. Furthermore, real-time interactions with the chatbot and the time machine in the VR chatbot provided a sense of authenticity, determining to what extent the learner could transfer what learners learned in the 3D virtual world to real life. Cai et al. (2020) study that conducted digital theatre with "the mechanisms of real-time spoken language evaluation and interactive switching of scenario & virtual costumes" found that students are more enthusiastic when performing in drama as they could wear digital theatre costumes and more willing to play the roles and read the scripts. Divekar et al. (2022) elucidated that the sense of immersion in VR chatbots provided students with culture absorption, incidental learning, learning out of need, and other benefits of actual in-country immersion programs, which could not be experienced in a physical classroom. Students learning using the VCA chatbot have identified human elements, which they perceive as a human-like partner or something between artefacts and humans. A strong tendency toward anthropomorphism indicates VCA's great potential as an interactive language partner in EFL contexts (Lee & Jeon, 2022).

Self-developed AI chatbots can bring compelling implications to language learners. According to Kim and Su (2024), the developed AI chatbot with the Danbee AI platform significantly improves students' willingness to communicate (WTC), reduces learning anxiety and enhances communication confidence. With the integration of XR, Divekar et al. (2022) also found that students feel fun and engaging when communicating in Chinese with a VR chatbot. The AI chatbot creates a more comfortable and less anxiety-ridden space for students to practice using the target language. As a result, it also improves students' WTC. The same situation was also found in Cai et al. (2020), where students are more willing to learn the learning materials and more seriously participate in the dramatic learning processes of digital theatre. Kohnke (2023) revealed that the AI chatbot can motivate students to complete out-of-class work and prepare for class more confidently. The same result was also found by Chiaráin and Chasaide (2016) that the design of ABAIR text-to-speech (TTS) in AI chatbot increases motivation and promotes positive attitudes towards learning the endangered language, such as the Irish language called Gaeilge. Haristiani and Rifa'I (2020) and Khlaisang and Sukavatee (2023) also support the idea that AI chatbots can improve students' motivation and enhance their learning environment. In addition, with additional portfolio management, Wu et al. (2024) SpokenBot manages to raise students' motivation and continuous self-learning. Setiawan et al. (2022) also elucidate that the Interactive English Learning chatbot can increase students' confidence in speaking English. The storybot designed by Barley et al. (2021) makes students more confident

using it by sending lots of messages.

Using One and More than One AI Chatbot Application

The application of one or multiple AI chatbots affects language teaching, primarily found in English language teaching. This phenomenon happened because most of the AI chatbots available in the market are made for English language learning. The current research findings are more about the implications of AI chatbots towards language proficiency, learning affective factors, and factors influencing students' perception of using AI chatbots. Most of the findings on language proficiency are on speaking skills, with five research studies: one on writing skills, reading proficiency, and learning performance, six on affective factors, and four on factors influencing the use of AI chatbots among students.

Firstly, for speaking performance, Kemelbekova et al. (2024) and Kim et al. (2021), in their research, discovered the same results in speaking proficiency even though they applied different AI chatbots. Kemelbekova et al. (2024) apply Text Cortex, a chatbot-driven AI system for vocabulary enrichment, Elsa Speak for voice recognition refinement and Grammarly for enhancing English grammar. Kim et al. (2021) apply several AI chatbot apps, such as Replika, Andy, and Google Assistant. The results showed that there is a significant improvement found between a group that applied AI chatbots and a group with conventional teaching in the classroom (Kemelbekova et al., 2024), as well as between low and intermediate levels of students (Kim et al., 2021) in oral communicative abilities, as evidenced by their performance in the two speaking tasks, namely reading a text aloud and answering questions. The same findings also found that there are no significant differences in terms of pronunciation. However, a notable difference is found in intonation and stress between the group that applied AI chatbots and the group with conventional teaching in the classroom (Kemelbekova et al., 2024) and between low and intermediate levels of students (Kim et al., 2021).

Moreover, Duong and Suppasetseree (2024) and Fathi et al. (2024) applied Andy English Chatbot. Duong and Suppasetseree (2024) indicate that students' English-speaking skills improved as they could speak English better using suitable hedging words, grammar structures, and vocabulary. Fathi et al. (2024) also support that AI-mediated interactive speaking activities improved EFL learners' speaking skills and willingness to communicate (WTC). Next, Lin and Mubarok (2021) applied the AI chatbot Replika in their mind map-guided flipped learning class. They found that the mind map-guided AI chatbot approach (MM-AI) promoted the students' English-speaking performances more than the conventional AI chatbot approach (C-AI). The significant result is that MM-AI can organise the interaction between robots and humans more than C-AI can. The AI chatbot also gives students positive attitudes and perceptions towards AI-mediated speaking instruction. Thus, it is proven that students believe using AI chatbots might stimulate an interaction cycle so they could practice the target language (Mohamed & Alian, 2023), enhancing their oral abilities.

Apart from that, for other language skills, such as writing, reading, and other language proficiency. For example, Alrajhi (2024), who applied a web-based AI pedagogical chatbot named Tutor Mike, revealed that AI chatbots are perceived as supportive of L2 practice and writing development. Nozhovnik et al. (2023), who applied chatbot-

driven e-classes on the Smart Sender platform, found that the application can improve students' English proficiency, especially in reading and language usage. Besides that, Çakmak (2022), who applied the AI chatbot Replika, found that student performance with the AI chatbot was significantly better than their face-to-face peer interactions. Annamalai et al. (2023) combination of multiple AI chatbots found that students who use chatbots gain proficiency in English and a stronger connection to language and the learning process.

Besides learning performance, AI chatbots available in the marketplace also impact learning affective factors, such as learning motivation, confidence, learning anxiety and WTC. For example, Alrajhi's (2024) research that applied Tutor Mike elucidates that the AI chatbot can enhance motivation and alleviate writing anxiety. By the way, most students enjoyed using AI chatbots and felt more comfortable when conversing with AI chatbots than a student partner or teacher. Hence, Annamalai et al., 2023 support that students who use AI chatbots gain confidence in using English. The same results were found by Mohamed and Alian (2023), which showed that the AI chatbots boosted students' enthusiasm and confidence, which ultimately helped them feel active and more comfortable. Tai and Chen (2020), who applied Google Assistant, also discovered similar results, where the AI chatbot provided a less threatening environment so that students could display a higher level of engagement, motivation, confidence and WTC in the target language. Fathi et al. (2024) also found that with the application of Andy English Chatbot, students improved their English-speaking skills and WTC. Nozhovnik et al. (2023) team, which applied chatbot-driven e-classes on the Smart Sender platform, discovered that the application could increase students' motivation.

However, there are different research findings found by El Shazly (2021) in his research that applied several AI web chatbot apps, such as Audrey, Charles, Cristal, and Mike, and Mondly, the written and oral communicative virtual partners. He revealed that students experienced certain anxiety levels at similar percentages pre- and post-intervention, and the anxiety levels were not alleviated when employing these AI chatbots; in fact, the Foreign Language Anxiety (FLA) increased slightly. This fact is supported by Çakmak (2022), who said there were negative perceptions and attitudes toward the chatbot interaction as students reported difficulties in being understood precisely, which might have contributed to higher anxiety in L2 speaking. Hence, Çakmak (2022) elucidates that although AI chatbot interaction is a novel way to provide speaking practice for students, interaction with a chatbot might not be a reliable way to lessen their anxiety with L2 speaking. Thus, more research needs to be done to fill this gap.

A few factors influence students' perception of using AI chatbots for language learning. Fryer et al. (2017) and (2019), who applied Cleverbot (Carpenter, n.d.) Chatbot, found that learning interests play a vital role in affecting students' learning with AI chatbots. Fryer et al. (2017) researched an application by comparing an AI chatbot with a human partner. They found that task interest had changed over time, with a decline in students' task interest with the AI chatbot but not the human partner. However, Fryer et al. (2019) discovered that although students reported having communication difficulties while using the AI chatbot, their experience of "learning more' with the chatbot was strongly connected to task interest. Hence, students' prior interest is strongly linked with language competency in the AI chatbot and is the best single predictor of future interest in AI chatbot conversation (Fryer et al., 2019).

Besides learning interest, according to Annamalai et al. (2023), Performance and Effort Expectancies contribute to the positive experience of using AI chatbots for language learning, hence serving as the Pull factors, which are positive factors that encourage one to accept the use of AI chatbots. The studies above have proved that applying one AI chatbot or more than one AI chatbot in language teaching and learning has provided positive implications for language proficiency. Hence, students have positive experiences using AI chatbots in their language learning. Meanwhile, Annamalai et al. (2023) also illustrate that Social Isolation contributed by the sense of robotic interaction, emotionlessness, and lack of flow in conversation serves as the Push Factor, which is considered the negative factors that prompt an individual to abandon the current choice in using AI chatbots for language learning. Hence, some research proves the facts. According to Alrajhi (2024), students' demotivating factors in using AI chatbot Tutor Mike are the lack of extended conversations, sensitivity to inaccurate language forms, and sporadic irrelevant responses. In contrast, Annamalai et al. (2023, 2023) in their two-research found that the push factors of the combination of more than one chatbot in interactive learning are lack of an emotional environment and personalised feedback and giving inaccurate or having issues in responsiveness in English language learning information. Also, Alm and Nkomo (2022) explain the students' frustration if the dialogues did not correspond to their learning goals, if they were excluded from using the bots because of technical or payment issues, or discontinuation of service.

In addition, perceived ease of use (PEOU) and perceived usefulness (PU) were crucial in using AI chatbots. This statement is agreed upon by students (Nozhovnik et al., 2023) and teacher candidates (Belda-Medina & Calvo-Ferrer, 2022). Nozhovnik et al. (2023) further illustrate that PEOU and PU of AI chatbots boost students' English language skills. However, Alrajhi (2024) only supports the usefulness of chatbot-mediated interaction, which has significantly more positive views from high-intermediate students. Since AI chatbot PU plays an important role, the factors that predict the PU are also essential. Therefore, Lorenzo et al. (2013), who applied a 3D environment with AI chatbots, discovered that the Communication and Collaboration Capabilities (CCC) variable is a significant antecedent to PU. Hence, the most essential determinant of AI chatbot adoption is the perceived value of cooperation, collaboration, communication and in-world assistance on the AI chatbot platform. Besides that, according to Belda-Medina and Calvo-Ferrer (2022), the gender of teacher candidates plays an essential factor in influencing the adoption of AI chatbots in the class as female candidates were more attentive to the use of inclusive design and language and more assertive about gender stereotyping.

The effects of AI chatbots on language performance and learning affective factors can be explained by Annamalai et al. (2023) based on three of the four behavioural engagements: constructive, active, and passive in Interactive, Constructive, Active, and Passive (ICAP) framework. The four distinct modes: Interactive, where students participate in discussions and collaborative activities; Constructive, which involves generating new knowledge through problem-solving and critical thinking; Active, focusing on practical, hands-on activities to apply concepts; and Passive, where students passively receive information through lectures or reading (Chi, 2009; Annamalai et al., 2023). These three modes expose students to language materials via meaningful activities in AI chatbots, contributing to their overall proficiency in the subject matter. However, according to Belda-Medina and Kokošková (2023), the Language Experience dimension (LEX), which includes elements such as Semantic Coherent Behaviour, Sentence Length and Complexity, and Speech Recognition and Synthesis, revealed that none

of the AI chatbots reached a moderate point of satisfaction among EFL teacher candidates. Hence, maybe most of the AI chatbots available in the marketplace do not fulfil the EFL teacher candidates' requirements, which need further investigation. Belda-Medina and Kokošková (2023) also suggest enhancements in adaptive user interfaces and incorporating social media and emerging technologies to simulate the human-student interaction and enrich the language learning experience. Besides that, this point also reveals the importance of self-developed AI chatbots, where teachers can build AI chatbots based on their students' personalised learning with social media integration. The self-developed AI chatbots should contain multimedia content, gamification and use of non-verbal language, which can become critical factors for learners' satisfaction, as indicated by the teacher candidates (Belda-Medina & Calvo-Ferrer, 2022).

Application of AI Chatbots with Integration into Social Media Applications or Platforms and Others

There are two research on the application of AI chatbots with Integration in Social Media Applications or platforms; where one is from the teachers' perspectives, who used it to teach English to secondary school students, while the other is based on high school students' English learning outcomes. Chuah and Kabilan (2021) researched and trained teachers to apply AI chatbots in the marketplace, such as Andy and Wordsworth, which integrated into Telegram and Facebook Messenger. The research found that teachers perceived using AI chatbots to give feedback to their students as very helpful, although some needed extra training on using them. Teachers thought AI chatbots could simulate an interaction cycle for students to practice the target language. Teachers also believed that chatbots augmented a more significant social presence, eventually creating an environment for their students to be active. Hence, teachers' points of view on AI chatbot applications matter to instructional designers and educational institutions. Besides that, Ye et al. (2022) applied Microsoft Xiaoying, an AI chatbot launched by Microsoft on WeChat, in high school students' English classes. They revealed that students' oral English grammar and pronunciation accuracy improved after practising the language with Xiaoying. Furthermore, most students use this AI chatbot positively as it improves oral accuracy.

Meanwhile, one research study that did not specify the types of AI chatbots and the language used to learn examines the correlations between students' understanding of AI chatbots, their engagement in using these technologies, their level of technological competence, and their achievements in language learning. They found a significant correlation between students' understanding of AI chatbots and their language learning achievements, with practical involvement mediating this relationship. They also discovered that technological proficiency was a significant moderator in influencing the link between knowledge and results. Hence, to get students to apply AI chatbots to enhance their language skills, students' practical involvement, technology proficiency, and understanding of AI chatbots play an essential role in their learning performance.

Conclusion

AI chatbots were developed over time to fulfil language education needs; they were created mainly to resolve the limitations of the physical language classes, such as (1) being limited to a particular time and place for learning, (2) being impractical for one-to-one student/teacher ratio; (3) language environment that cannot supply enough

chance of authentic talking (Jia, 2009; Fryer & Carpenter, 2006). Hence, Fryer and Carpenter (2006) also added some aspects of AI chatbots that enhance language learning: (1) chatbot can make learners motivate in learning; (2) learners feel more relaxed when chatting with a bot rather than with an actual human partner; (3) chatbot can repeat the same material for several times without being bored; (5) many chatbots use texts and speech modality for conversation, which are helpful to enhance the learners' communication skills, such as listening, speaking, reading and writing (Shawar, 2017). Hence, the study's conclusion relates to (1) the state-of-the-art AI chatbots in language education between 2006 and 2024, (2) the implications of different AI chatbots' influence on teachers' and students' perspectives on language teaching and learning and students' learning outcomes and (3) recommendations for future research.

AI chatbot was first designed to interact with and entertain native speakers (Fryer & Carpenter, 2006). AI chatbots continue to develop to cater to the need for ESL and EFL as their learning partner in facilitating their learning of English with the medium of explanation being their native language. The development also started in foreign and endangered languages like the Irish. Some teaching and learning can apply to ready-made AI chatbots found in the marketplace, and some may need educators or designers to develop AI chatbots to fulfil teaching and learning needs. These needs increased tremendously from 2020 until 2022 as the world was affected by the COVID-19 lockdown that caused all teaching and learning to be online. Hence, applications of self-developed AI chatbots increased as educators could self-design them using Google's Dialogflow, BlenderBot, or any other available AI chatbot builder. Some self-developed AI chatbots were integrated with social media applications or platforms. Designing an AI chatbot also taught learners American Sign Language words and phrases. At the end of 2022, Large Language Model (LLM) or Generative AI (Gen AI) chatbots, such as ChatGPT and Bing's Chat, were introduced into the world. The emergence of the new AI chatbots enables users to provide questions with prompts and get answers faster. Hence, the LLMs AI chatbots started to play a role in language teaching and learning as language learning tools and how they can facilitate language teaching and learning for teachers and students.

AI chatbots have implications for language learning, especially for learning outcomes, learning affective factors, and learning factors from students' and teachers' perspectives. Firstly, for LLMs, AI chatbots like ChatGPT and Bing's Chat. The applications of LLMs AI chatbots have favourable implications for language learning outcomes with teacher guidance. Hence, teachers' perspectives on LLMs AI chatbots matter. The study discovered that teachers tend to be worried about AI-based academic dishonesty and students' over-reliance on using Gen AI chatbots, and most of these teachers have not previously applied LLMs AI chatbots in their teaching and learning. Therefore, teachers who have used LLMs show higher positivity and lower negativity beliefs than those who have never used LLMs before. Students have different perspectives than teachers as they exhibit critical judgment in evaluating the quality of ideas and outputs generated by ChatGPT and the ability to modify prompts to maximise learning benefits. For students to make essential decisions about the quality of ideas provided by ChatGPT, they need to have the ability to (1) understand the nature of ChatGPT, their interactions and experimentation with ChatGPT, and the guidance provided by their teachers.

Second, self-developed AI chatbots with and without integration into social media applications or platforms impacted language learning performance, self-engagement, and learning affective factors. Both self-developed AI

chatbots affect language learning proficiency positively, except there is no research on the application of self-developed AI chatbots with integrated social media applications or platforms for writing skills. The self-developed AI chatbots provide human-like interaction for learning engagement, enhancing students' sense of engagement. Hence, some research has proven that students tend to have longer engagement time, enrich communicative strategies, and produce more language outputs as they engage with AI chatbot applications. Students can experience real-world situations in the target language when a virtual environment and voiced-controlled conversational agents (VCAs) are integrated into self-developed AI chatbots. These kinds of experiences provide students with a sense of immersion and authenticity with culture absorption, incidental learning, learning out of need, and other benefits of actual in-country immersion programs, which could not be experienced in a physical classroom (Wang et al., 2018; Divekar et al., 2022). Both self-developed AI chatbots facilitate students in language learning, which significantly reduces learning anxiety and improves learning motivation, learning confidence and willingness to communicate. This phenomenon happened because AI chatbots create a more comfortable and less anxiety-ridden space for students to practice using the target language.

Next, applying single or multiple AI chatbots has implications for language learning performance, learning affective factors, and factors influencing students' and teachers' use of AI chatbots. Research has proven that the application of one or more than one AI chatbot has positive impacts on English language learning performance and learning affective factors, except El Shazly (2021), who applied several AI web chatbot apps, such as Audrey, Charles, Cristal, and Mike, and Mondly. He revealed that students experienced certain anxiety levels at similar percentages pre- and post-intervention, and the anxiety levels were not alleviated when employing these AI chatbots but have increased. This situation can be due to difficulties in being understood precisely, and the interaction with AI chatbots might not be a reliable way to lessen students' L2 speaking anxiety (Çakmak, 2022), which needs to be further verified. Factors influencing students' learning performance in AI chatbot applications are learning interests and AI chatbot language proficiency (Fryer et al., 2017, 2019). Besides that, Annamalai et al. (2023) added the Pull factors contributing to the positive experience of using Chatbots, like Performance and Effort Expectancies, and Pull factors that cause negativity in using Chatbots, such as sense of robotic interaction, emotionlessness, and lack of flow in conversation that contribute to Social Isolation. Meanwhile, the PEOU and PU are the main factors influencing teachers and students to use AI chatbots in language teaching and learning. Gender also plays an essential role, as female candidates were more attentive to using inclusive design and language and more assertive about gender stereotyping (Belda-Medina & Calvo-Ferrer, 2022). Based on Annamalai et al. (2023), three of the four behavioural engagements: constructive, active, and passive in the Interactive, Constructive, Active, and Passive (ICAP) framework expose students to language materials via meaningful activities in AI chatbots, contributing to their overall proficiency in the subject matter. However, Belda-Medina and Kokošková (2023) revealed that none of the AI chatbots reached a moderate point of satisfaction among EFL teacher candidates, especially in the Language Experience dimension (LEX). Since the ready-made AI chatbots in the market do not fulfil the teachers' needs, they suggest enhancements in adaptive user interfaces and incorporating social media and emerging technologies to simulate the human-student interaction and enrich the language learning experience. Hence, this points out the need for a self-developed AI chatbot to integrate with social media applications, platforms, or any virtual environment containing multimedia content, gamification and non-verbal language.

Most studies only focus on particular AI chatbots with particular backgrounds and a limited number of teachers and students, which could not provide generalised results. The authors also found that future studies will focus on ChatGPT, English language teaching, language learning model, technology acceptance, second language writing and Dialogflow. Hence, for future studies, it is recommended that (1) extend the studies toward students/teachers from various regions, language proficiency levels, and communities with different cultural backgrounds, (2) longitudinal research could be employed to see if there is any novelty effect or other changes in the learning outcomes, affective gains and factors influence the use of the AI chatbots over an extended period, (3) focus on developing strategies, language learning model and process, teaching approaches or methods, assistance from teachers and peers and guidelines for integrating AI chatbots, especially with LLMs AI chatbots into curriculum effectively, (4) effects of learning with self-developed AI chatbots or LLMs AI chatbots that are integrated with more intelligence, realistic agents capable of performing several expression, gestures and movements or more additional games, quizzes, and more multimedia elements in enhancing language learning, (5) factors influence teachers and students in acceptance the use of AI chatbots.

References

Abdul-Kader, S.A.; and Woods, J. (2015). Survey on chatbot design techniques in speech conversation systems. *International Journal of Advanced Computer Science and Applications, 6(7),* 72-80.

Alm, A., & Nkomo, L. M. (2022). Chatbot experiences of informal language learners: A sentiment analysis. In *Research anthology on implementing sentiment analysis across multiple disciplines* (pp. 933-948). Igi Global.

Annamalai, N., Ab Rashid, R., Hashmi, U. M., Mohamed, M., Alqaryouti, M. H., & Sadeq, A. E. (2023). Using chatbots for English language learning in higher education. *Computers and Education: Artificial Intelligence, 5*, 100153. https://doi.org/10.1016/j.caeai.2023.100153

Annamalai, N., Eltahir, M. E., Zyoud, S. H., Soundrarajan, D., Zakarneh, B., & Al Salhi, N. R. (2023). Exploring English language learning via Chabot: A case study from a self-determination theory perspective. *Computers and Education: Artificial Intelligence, 5*, 100148. https://doi.org/10.1016/j.caeai.2023.100148

Annamalai, N., Zakarneh, B., Mangaleswaran, S., & Zyoud, S. H. (2023). Exploration of undergraduate students lived experiences of using a chatbot for English language learning. *International Journal of English Language and Literature Studies, 12(3),* 254-268. DOI: 10.55493/5019.v12i3.4876

Assim S. Alrajhi (2024) Artificial intelligence pedagogical chatbots as L2 conversational agents, *Cogent Education, 11:1*, 2327789, https://doi.org/10.1080/2331186X.2024.2327789

Azwary, F.; Indriani, F.; and Nugrahadi, D.T. (2016). Question answering system berbasis artificial intelligence markup language sebagai media informasi. *Klik-Kumpulan Jurnal Ilmu Komputer, 4(1*), 48-60.

Bailey, D., & Almusharraf, N. (2021, April). Investigating the effect of chatbot-to-user questions and directives on student participation. In 2021 *1st International Conference on Artificial Intelligence and Data Analytics (CAIDA)* (pp. 85-90). IEEE. DOI: 10.1109/CAIDA51941.2021.9425208

Bailey, D., Southam, A., & Costley, J. (2021). Digital storytelling with chatbots: Mapping L2 participation and perception patterns. *Interactive Technology and Smart Education, 18(1),* 85-103.

http://dx.doi.org/10.1108/ITSE-08-2020-0170

Belda-Medina, J., & Calvo-Ferrer, J. R. (2022). Using chatbots as AI conversational partners in language learning. *Applied Sciences, 12(17),* 8427. https://doi.org/10.3390/app12178427

Belda-Medina, J., & Kokošková, V. (2023). Integrating chatbots in education: insights from the Chatbot-Human Interaction Satisfaction Model (CHISM). *International Journal of Educational Technology in Higher Education, 20(1),* 62. https://doi.org/10.1186/s41239

Berns, A., Mota, J. M., Ruiz-Rube, I., & Dodero, J. M. (2018, October). Exploring the potential of a 360 video application for foreign language learning. In *Proceedings of The Sixth International Conference On Technological Ecosystems For Enhancing Multiculturality* (pp. 776-780). https://doi.org/10.1145/3284179.3284309.

Cai, M. Y., Wang, J. Y., Chen, G. D., Wang, J. H., & Yang, S. H. (2020, July). A digital reality theater with the mechanisms of real-time spoken language evaluation and interactive switching of scenario & virtual costumes: Effects on motivation and learning performance. In 2020 IEEE *20th International Conference on Advanced Learning Technologies (ICALT)* (pp. 295-299). IEEE. DOI 10.1109/ICALT49669.2020.00096

Çakmak, F. (2022). Chatbot-Human Interaction and Its Effects on EFL Students' L2 Speaking Performance and Anxiety. *Novitas-ROYAL (Research on Youth and Language), 16(2),* 113-131.

Chen, H. L., Vicki Widarso, G., & Sutrisno, H. (2020). A chatbot for learning Chinese: Learning achievement and technology acceptance. *Journal of Educational Computing Research, 58(6),* 1161-1189. http://dx.doi.org/10.1177/0735633120929622

Chi, M. T. (2009). Active-constructive-interactive: A conceptual framework for differentiating learning activities. Topics in Cognitive Science, 1(1), 73-105. https://doi.org/10.1111/j.1756-8765.2008.01005.x

Chiaráin, N. N., & Chasaide, A. N. (2016, May). Chatbot technology with synthetic voices in the acquisition of an endangered language: motivation, development and evaluation of a platform for Irish. In Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16) (pp. 3429-3435).

Cong-Lem, N., Tran, T. N., & Nguyen, T. T. (2024). Academic Integrity In The Age of Generative AI: Perceptions and Responses of Vietnamese EFL Teachers. *Teaching English with Technology, 24(1),* 28–47. https://doi.org/10.56297/FSYB3031/MXNB756

Coniam, D. (2008). Evaluating the language resources of chatbots for their potential in English as a second language. *ReCALL, 20(1),* 98-116.

De Iacovo, V., Palena, M., & Romano, A. (2021). Evaluating prosodic cues in Italian: the use of a Telegram chatbot as a CALL tool for Italian L2 learners. In *L'individualità del Parlante Nelle Scienze Fonetiche: Applicazioni Tecnologiche E Forensi* (pp. 283-298). https://dx.doi.org/10.17469/O2108AISV000015

Divekar, R. R., Lepp, H., Chopade, P., Albin, A., Brenner, D., & Ramanarayanan, V. (2021, July). Conversational agents in language education: where they fit and their research challenges. *In International Conference on Human-Computer Interaction* (pp. 272-279). Cham: Springer International Publishing. https://link.springer.com/chapter/10.1007/978-3-030-90179-0_35

Dehghani, H., & Mashhadi, A. (2024). Exploring Iranian english as a foreign language teachers' acceptance of ChatGPT in english language teaching: Extending the technology acceptance model. *Education and*

Information Technologies, 1-22. https://doi.org/10.1007/s10639-024-12660-9

Donthu, N., Kumar, S., Mukherjee, D., Pandey, N., & Lim, W.M. (2021). How to conduct a bibliometric analysis: An overview and guidelines. *Journal of Business Research, 133,* 285–296. https://doi.org/10.1016/j.jbusres.2021.04.070

Duong, T., & Suppasetseree, S. (2024). The Effects of an Artificial Intelligence Voice Chatbot on Improving Vietnamese Undergraduate Students' English Speaking Skills. *International Journal of Learning, Teaching and Educational Research, 23(3),* 293-321. https://doi.org/10.26803/ijlter.23.3.15

El Shazly, R. (2021). Effects of artificial intelligence on English speaking anxiety and speaking performance: A case study. *Expert Systems, 38(3),* e12667. https://doi.org/10.1111/exsy.12667

Escalante, J., Pack, A., & Barrett, A. (2023). AI-generated feedback on writing: insights into efficacy and ENL student preference. *International Journal of Educational Technology in Higher Education, 20(1),* 57. https://doi.org/10.1186/s41239-023-00425-2

Fathi, J., Rahimi, M., & Derakhshan, A. (2024). Improving EFL learners' speaking skills and willingness to communicate via artificial intelligence-mediated interactions. *System, 121,* 103254. https://doi.org/10.1016/j.system.2024.103254

Fryer, L. K., Ainley, M., Thompson, A., Gibson, A., & Sherlock, Z. (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. http://dx.doi.org/10.1016/j.chb.2017.05.045

Fryer, L., & Carpenter, R. (2006). Bots as Language Learning Tools. *Language Learning & Technology, 10(3),* 8-14.

Fryer, L. K., Nakao, K., & Thompson, A. (2019). Chatbot learning partners: Connecting learning experiences, interest and competence. *Computers in human Behavior, 93,* 279-289. https://doi.org/10.1016/j.chb.2018.12.023

Gan, L. (2024). Application of Fuzzy Neural Network in Teaching Spoken English for Tourism. *Informatica, 48(5).* https://doi.org/10.31449/inf.v48i5.5300

Ghafouri, M., Hassaskhah, J., & Mahdavi-Zafarghandi, A. (2024). From virtual assistant to writing mentor: Exploring the impact of a ChatGPT-based writing instruction protocol on EFL teachers' self-efficacy and learners' writing skill. *Language Teaching Research*, 1-23. https://doi.org/10.1177/13621688241239764

Goyal, P., Pandey, S., & Jain, K. (2018). Developing a chatbot. Available from *Deep Learning for Natural Language Processing, Apress, Berkeley, CA* https://doi.org/10.1007/9781-4842-3685-7_4

Guangxiang Leon Liu, Ron Darvin & Chaojun Ma (2024, Feb 02): Exploring AI-mediated informal digital learning of English (AI-IDLE): a mixed-method investigation of Chinese EFL learners' AI adoption and experiences, *Computer Assisted Language Learning*, DOI: 10.1080/09588221.2024.2310288

Guo, K., Li, Y., Li, Y., & Chu, S. K. W. (2024). Understanding EFL students' chatbot-assisted argumentative writing: An activity theory perspective. *Education and Information Technologies, 29(1),* 1-20. https://doi.org/10.1007/s10639-023-12230-5

Haristiani, N., Dewanty, V. L., & Rifai, M. M. (2022). Autonomous learning through chatbot-based application utilization to enhance basic Japanese competence of vocational high school students. *Journal of Technical Education and Training, 14(2),* 143-155. https://doi.org/10.30880/jtet.2022.14.02.013

Haristiani, N., & Rifa'i, M. M. (2020). Combining chatbot and social media: Enhancing personal learning environment (PLE) in language learning. *Indonesian Journal of Science and Technology, 5(3*), 487-506. http://dx.doi.org/10.17509/ijost.v5i3.2868

Haristiani, N., & Rifai, M. M. (2021). Chatbot-based application development and implementation as an autonomous language learning medium. Indonesian Journal of Science and Technology, 6(3), 561-576. https://doi.org/10.17509/ijost.v

Huang, J., Zhou, M., & Yang, D. (2007). Extracting chatbot knowledge from online discussion forums. *International Joint Conferences on Artificial Intelligence, 7,* 423-428.

Jain, M., Kumar, P., Kota, R. and Patel, S.N. (2018) 'Evaluating and informing the design of chatbots', In Proceedings of the 2018 on *Designing Interactive Systems Conference 2018*, pp.895–906, Hong Kong, China: ACM. doi:10.1145/3196709.3196735

Jaleniauskienė, E., Lisaitė, D., & Daniusevičiūtė-Brazaitė, L. (2023). Artificial intelligence in language education: a bibliometric analysis. *Sustainable Multilingualism, 23(1),* 159-194.

Jia, J. (2009). CSIEC: A computer assisted English learning chatbot based on textual knowledge and reasoning. *Knowledge-based systems, 22(4),* 249-255.

Johnson, C., Urazov, M., & Zanoli, E. (2022). Escapeling: a gamified, AI-supported chatbot for collaborative language practice. In Innovations in Learning and Technology for the Workplace and Higher Education: *Proceedings of 'The Learning Ideas Conference'2021* (pp. 141-148). Springer International Publishing.

Juan Wu, Yuxin Li, Jianrong Zhou & Shiya Chen (2024, Feb 19): The impact of intelligent personal assistants on Mandarin second language learners: interaction process, acquisition of listening and speaking ability, *Computer Assisted Language Learning*, DOI: 10.1080/09588221.2024.2317849

Karataş, F., Abedi, F. Y., Ozek Gunyel, F., Karadeniz, D., & Kuzgun, Y. (2024). Incorporating AI in foreign language education: An investigation into ChatGPT's effect on foreign language learners. *Education and Information Technologies*, 1-24. https://doi.org/10.1007/s10639-024-12574-6

Kemelbekova, Zada & Degtyareva, Xeniya & Yessenaman, Saule & Ismailova, Dariga. (2024). AI in teaching English as a foreign language: Effectiveness and prospects in Kazakh higher education. XLinguae. 17. 69-83. DOI: 10.18355/XL.2024.17.01.05

Kerly, A., Hall, P. and Bull, S. (2007) 'Bringing chatbots into education: Towards natural language negotiation of open learner models', *Knowledge-Based Systems*, Vol. 20, No. 2, pp.177–185.

Kharis, M., Schön, S., Hidayat, E., Ardiansyah, R., & Ebner, M. (2022). Mobile Gramabot: Development of a chatbot app for interactive German grammar learning. *International Journal of Emerging Technologies in Learning (iJET), 17(14)*, 52-63. https://doi.org/10.3991/ijet.v17i14.31323

Khlaisang, J., & Sukavatee, P. (2023). Mobile-Assisted Language Learning to Support English Language Communication among Higher Education Learners in Thailand. *Electronic Journal of e-Learning, 21(3),* 234-247.

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. http://hdl.handle.net/10125/73463

Kim, H. S., Cha, Y., & Kim, N. Y. (2021). Effects of AI chatbots on EFL students' communication skills. *Korean Journal of English Language and Linguistics (21)*, 712-734. DOI: 10.15738/kjell.21..202108.712

Kim, A., & Su, Y. (2024). How implementing an AI chatbot impacts Korean as a foreign language learners'

willingness to communicate in Korean. *System, 103256*. https://doi.org/10.1016/j.system.2024.103256

Kohnke, L. (2022). 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.

Kohnke, L. (2023). A pedagogical chatbot: A supplemental language learning tool. *RELC Journal, 54(3),* 828-838. DOI: 10.1177/00336882211067054

Kostikova, I., Holubnycha, L., Besarab, T., Moshynska, O., Moroz, T., Shamaieva, I. (2024). ChatGPT for Professional English Course Development. *International Journal of Interactive Mobile Technologies (iJIM), 18(2)*, pp. 68–81. https://doi.org/10.3991/ijim.v18i02.46623

Kwon, S. K., Shin, D., & Lee, Y. (2023). The application of chatbot as an L2 writing practice tool. *Language Learning & Technology, 27(1)*, 1–19. https://doi.org/10125/73541

Lee, H., Hsia, C. C., Tsoy, A., Choi, S., Hou, H., & Ni, S. (2023, September). VisionARy: Exploratory research on Contextual Language Learning using AR glasses with ChatGPT. In *Proceedings of the 15th Biannual Conference of the Italian SIGCHI* Chapter (pp. 1-6).

Lee, J. H., Yang, H., Shin, D., & Kim, H. (2020). Chatbots. *ELT Journal, 74(3),* 338–344. https://doi.org/10.1093/elt/ccaa035

Lee, K. A., Lim, S. B., & Nagarajan, S. N. (2023). A Study of the Effectiveness of English Speaking of Teachable Agent using AI Chatbot. In *ICAART (1)* (pp. 308-314). DOI: 10.5220/0011730300003393

Lee, S., Jeon, J., & Choe, H. (2024). Enhancing Pre-Service Teachers' Global Englishes Awareness with Technology: A Focus on AI Chatbots in 3D Metaverse Environments. *TESOL Quarterly*. DOI: 10.1002/tesq.3300

Li, X., Li, B., & Cho, S. J. (2023). Empowering Chinese Language Learners from Low-Income Families to Improve Their Chinese Writing with ChatGPT's Assistance Afterschool. *Languages, 8(4),* 238. https://doi.org/10.3390/languages8040238

Liang, K. H., Davidson, S., Yuan, X., Panditharatne, S., Chen, C. Y., Shea, R., ... & Fryer, L. (2023, July). ChatBack: Investigating methods of providing grammatical error feedback in a GUI-based language learning chatbot. In Proceedings of the *18th Workshop on Innovative Use of NLP for Building Educational Applications (BEA 2023)* (pp. 83-99).

Lin, C.-J., & Mubarok, H. (2021). Learning Analytics for Investigating the Mind Map-Guided AI Chatbot Approach in an EFL Flipped Speaking Classroom. *Educational Technology & Society, 24 (4)*, 16–35.

Liu, G., & Ma, C. (2023, in press). Measuring EFL learners' use of ChatGPT in informal digital learning of English based on the technology acceptance model. *Innovation in Language Learning and Teaching*, 1-18.

Mohamed, S. S.A., & Alian, E. M. I. (2023). Students' Attitudes toward Using Chatbot in EFL Learning. *Arab World English Journal, 14 (3)* 15-27. DOI: https://dx.doi.org/10.24093/awej/vol14no3.2

Mu, J., & Sarkar, A. (2019, May). Do we need natural language? Exploring restricted language interfaces for complex domains. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems* (pp. 1-6).

Nozhovnik, O., Harbuza, T., Teslenko, N., Okhrimenko, O., Zalizniuk, V., & Durdas, A. (2023). Chatbot gamified

and automated management of L2 learning process using smart sender platform. International Journal of Educational Methodology, 9(3), 603-618. https://doi.org/10.12973/ijem.9.3.603

Oralbayeva, N., Shakerimov, A., Sarmonov, S., Kantoreyeva, K., Dadebayeva, F., Serkali, N., & Sandygulova, A. (2022, March). K-qbot: Language learning chatbot based on reinforcement learning. In *2022 17th ACM/IEEE International Conference on Human-Robot Interaction (HRI)* (pp. 963-967). IEEE.

Paudyal, P., Banerjee, A., & Gupta, S. (2020, March). On evaluating the effects of feedback for sign language learning using explainable AI. In *Proceedings of the 25th International Conference on Intelligent User Interfaces Companion* (pp. 83-84). https://doi.org/10.1145/3379336.3381469

Pereira, J. and Díaz, Ó. (2018) 'Chatbot dimensions that matter: Lessons from the trenches', In *International Conference on Web Engineering*, Springer, Cáceres, Spain, pp.129–135, DOI: 10.1007/978-3-319-91662-0_9.

Pham, X. L., Pham, T., Nguyen, Q. M., Nguyen, T. H., & Cao, T. T. H. (2018, November). Chatbot as an intelligent personal assistant for mobile language learning. In *Proceedings of the 2018 2nd International Conference on Education and E-Learning* (pp. 16-21). https://doi.org/10.1145/3291078.3291115

Polyzi, P., & Moussiades, L. (2023). An artificial vocabulary learning assistant. *Education and Information Technologies, 28(12),* 16431-16455. https://doi.org/10.1007/s10639-023-11810-9

Poseletska, K., Kyrychenko, S., Vlasenko, O., Koval, I., Potiuk, I. & Shpenyk, S. (2023). Scenarios for the Use of Chatbots in Teaching a Foreign Language in the Higher Educational Institution (HEI). *Revista Românească pentru Educaţie Multidimensională, 15(3),* 347-359. https://doi.org/10.18662/rrem/15.3/770

Punar Özçelik, N., & Yangın Ekşi, G. (2024). Cultivating writing skills: the role of ChatGPT as a learning assistant—a case study. *Smart Learning Environments, 11(1),* 10. https://doi.org/10.1186/s40561-024-00296-8

Qianqian Cai, Yupeng Lin & Zhonggen Yu (2023, Oct 15): Factors Influencing Learner Attitudes Towards ChatGPT-Assisted Language Learning in Higher Education, *International Journal of Human–Computer Interaction*, DOI: 10.1080/10447318.2023.2261725

Qian, K., Shea, R., Li, Y., Fryer, L. K., & Yu, Z. (2023, June). User Adaptive Language Learning Chatbots with a Curriculum. In *International Conference on Artificial Intelligence in Educatio*n (pp. 308-313). Cham: Springer Nature Switzerland.

Qin, R., & Yu, Z. (2022). A Bibliometric Analysis of Students' Collaborative Learning and Online Social Presence via Tencent Meeting and WeChat. *International Journal of Online Pedagogy and Course Design (IJOPCD),* 12(4), 1-21.

Qu, K., & Wu, X. (2024). ChatGPT as a CALL tool in language education: A study of hedonic motivation adoption models in English learning environments. *Education and Information Technologies*, 1-33. https://doi.org/10.1007/s10639-024-12598-y

Rodosthenous, C., Lyding, V., Sangati, F., König, A., ul Hassan, U., Nicolas, L., ... & Aparaschivei, L. (2020). Using crowdsourced exercises for vocabulary training to expand conceptnet. In 1*2th International Conference on Language Resources and Evaluation (LREC 2020*).

Ruan, S., Jiang, L., Xu, Q., Liu, Z., Davis, G. M., Brunskill, E., & Landay, J. A. (2021, April). Englishbot: An ai-powered conversational system for second language learning. In *Proceedings of the 26th International*

Conference on Intelligent User Interfaces (pp. 434-444). https://doi.org/10.1145/3397481.3450648

Ruan, S., Willis, A., Xu, Q., Davis, G. M., Jiang, L., Brunskill, E., & Landay, J. A. (2019, June). Bookbuddy: Turning digital materials into interactive foreign language lessons through a voice chatbot. In *Proceedings of the sixth (2019) ACM conference on learning@ scal*e (pp. 1-4). https://doi.org/10.1145/3330430.3333643

Seongyong Lee & Jaeho Jeon (2022): Visualizing a disembodied agent: young EFL learners' perceptions of voice-controlled conversational agents as language partners, *Computer Assisted Language Learning*, DOI: 10.1080/09588221.2022.2067182

Setiawan, P. R., Nasution, A. H., Hanafia, A., Arta, Y., Ramadhan, R. A., & Kadir, E. A. (2022, October). Interactive English teaching and learning based on mobile application. In *2022 3rd International Conference on Electrical Engineering and Informatics (ICon EEI)* (pp. 181-185). IEEE. DOI: 10.1109/IConEEI55709.2022.9972321

Shawar, B. A., & Atwell, E. (2007). Chatbots: are they really useful?. *Journal for Language Technology and Computational Linguistics, 22(1),* 29-49.

Shevat, A. (2017). *Designing bots: Creating conversational experience*s. Sebastopol, California, United States of America: O'Reilly Media, Inc.

Shin, D., Lee, J. H., & Noh, W. I. (2024). Realizing Corrective Feedback in Task-Based Chatbots Engineered for Second Language Learning. *RELC Journal, 00336882231221902*. DOI: 10.1177/00336882231221902

Tam, A. (2023, June 22). *What are large language models. Machine Learning Mastery*. https://machinelearningmastery.com/what-are-large-language-models/#:~:text=There%20are%20multiple%20large%20language,language%20and%20can%20generate%20text.

van Eck, N.J., & Waltman, L. (2010). Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics, 84(2),* 523–538. doi:10.1007/s11192-009-0146-3

Van Eck, N. J., & Waltman, L. (2020). *VOSviewer manual: Manual for VOSviewer version 1.6.16. Vol. 1:* 1–53. Leiden: Universitat Leiden. 1. https://www.vosviewer.com/documentation/Manual_VOSviewer_1.6.16.pdf

Wang, Y. F., & Petrina, S. (2013). Using learning analytics to understand the design of an intelligent language tutor–Chatbot Lucy. *International Journal of Advanced Computer Science and Applications, 4(11),* 124–131. https://doi.org/10.14569/IJACSA.2013.041117

Weizebaum, J. (1966). ELIZA--A computer program for the study of natural language communication between man and Machine [Electronic Version]. *Communications of the ACM, 9*. http://i5.nyu.edu/~mm64/x52.9265/january1966.html.

Wu, C. H., Lam, H. Y., Kong, A., & Wong, W. L. H. (2023). The design and evaluation of a digital learning-based English chatbot as an online self-learning method. *International Journal of Engineering Business Management, 15*, 18479790231176372. DOI: 10.1177/18479790231176372

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. https://doi.org/10.3390/languages8030212

Yang, H., Kim, H., Lee, J. H., & Shin, D. (2022). Implementation of an AI chatbot as an English conversation partner in EFL speaking classes. *ReCALL, 34(3),* 327-343.

Yang Gao, Qikai Wang & Xiaochen Wang (2024, Jan 21): Exploring EFL university teachers' beliefs in

integrating ChatGPT and other large language models in language education: a study in China, *Asia Pacific Journal of Education*, DOI: 10.1080/02188791.2024.230517

Ye, Y., Deng, J., Liang, Q., & Liu, X. (2022). Using a Smartphone-Based Chatbot in EFL Learners' Oral Tasks. *International Journal of Mobile and Blended Learning (IJMBL), 14(1)*, 1-17. DOI: 10.4018/IJMBL.299405

Zhang, Z., & Huang, X. (2024). The impact of chatbots based on large language models on second language vocabulary acquisition. *Heliyon, 10(3)*, 1-13. https://doi.org/10.1016/j.heliyon.2024.e25370

Author Information

Hui Wen Chua
https://orcid.org/0000-0002-5181-8246
Faculty of Language Studies and Human Development, Universiti Malaysia Kelantan.
Kelantan, Malaysia
& School of Distance Education, Universiti Sains Malaysia.
Penang, Malaysia

Nagaletchimee Annamalai
https://orcid.org/0000-0002-8600-5624
School of Distance Education, Universiti Sains Malaysia.
Penang, Malaysia
Contact e-mail: *naga@usm.my*

Appendix. Different Types of Chatbots in Language Learning

No.	Authors	Type of AI Chatbot						Platform / Application	Language					Subject				
		C	S	W	O	M	I		E	NM	FL			T	UG	MH	P	NM
											G	M	J OL					
1.	Cai et al., 2023	V						ChatGPT		V					V			
2.	Cong-Lem et al., 2024	V						ChatGPT	V					V	V			
3.	Escalante et al., 2023	V						ChatGPT (GPT-4)	V						V			
4.	Gao et al., 2024	V						ChatGPT	V					V	V			
5.	Karataş et al., 2024	V						ChatGPT	V						V			
6.	Kostikova et al., 2024	V						ChatGPT	V						V			
7.	Lee et al., 2023	V						VisionARy, integration ChatGPT into AR glasses	V									V
8.	Ghafouri et al., 2024	V						ChatGPT	V					V	V			
9.	Li et al., 2023	V						ChatGPT 3.5				V				V		
10.	Liu et al., 2024	V						Large language model platforms (e.g. ChatGPT, Bing Chat	V						V			
11.	Punar Özçelik & Yangın Ekşi, 2024	V						ChatGPT	V						V			
12.	Qu & Wu, 2024	V						ChatGPT	V						V			
13.	Shaikh et al., 2023	V						ChatGPT	V									V
14.	Xiao & Zhi, 2023	V						ChatGPT		V					V			
15.	Zhang & Huang, 2024	V						Self-developed Large Language Model (LLM)	V							V		
16.	Dehghani & Mashhadi, 2024	V						ChatGPT	V					V		V		
17.	Liu & Ma, 2024	V						ChatGPT	V						V	V		
18.	Lee, et al, 2023		V					Teachable Agent: Odinga Agent	V					V		V		
19.	Berns et al., 2018		V					Develop a chatbot application			V				V			

No.	Authors	Type of AI Chatbot						Platform / Application	Language						Subject				
		C	S	W	O	M	I		E	NM	FL				T	UG	MH	P	NM
											G	M	J	OL					
								named 'Let's date!' by using an authoring tool named Visual Environment for Designing Interactive Learning Scenarios (VEDILS)											
20.	Cai et al., 2020							Microsoft speech recognition, speech synthesis, and Language Understanding Intelligent Service (LUIS) cloud services are used to construct an AI chatbot. (a realistic speech mission-based digital learning theatre)					V			V			
21.	Divekar et al., 2022		V					Combination of Artificial Intelligence (AI) and Extended Reality (XR) to create The Cognitive Immersive Language Learning Environment (CILLE)				V				V			
22.	Gan, 2024		V					Development of an interactive language learning platform using a fuzzy neural network (FNN)	V							V			
23.	Guo et al., 2024		V					Chatbot,	V							V			

No.	Authors	Type of AI Chatbot						Platform / Application	Language						Subject				
		C	S	W	O	M	I		E	NM	FL				T	UG	MH	P	NM
											G	M	J	OL					
								Argumate											
24.	Jia, 2009		V					CSIEC (Computer Simulation in Educational Communication) system	V							V	V		
25.	Kharis et al., 2022		V					Mobile chatbot, Gramabot			V					V			
26.	Kim & Su, 2024		V					AI chatbot developed with Danbee AI platform.						V		V			
27.	Kim et al., 2022		V					Chatbot Ellie	V								V		
28.	Kohnke, 2023		V					AI chatbot developed with Google's Dialogflow.	V						V	V			
29.	Kwon et al., 2023		V					AI chatbot was developed with Google's Dialogflow by encoding expressions from an elementary English textbook.	V								V		
30.	Lee et al., 2024		V					AI Chatbot and Metaverse	V						V				
31.	Liang et al., 2023		V					Chatback, an AI chatbot that was developed with Blenderbot3 3B	V										V
32.	Paudyal et al., 2020		V					SignGuru, a chatbot-based AI tutor						V					V
33.	Polyzi, & Moussiades, 2023		V					Self-developed AI chatbot	V										V
34.	Qian et al., 2023		V					Ai chatbot was developed with Blenderbot3 3B	V								V		
35.	Ruan et al., 2021		V					AI chatbot named	V							V			

No.	Authors	Type of AI Chatbot						Platform / Application	Language						Subject				
		C	S	W	O	M	I		E	NM	FL				T	UG	MH	P	NM
											G	M	J	OL					
								EnglishBot											
36.	Ruan et al., 2019		V					Bookbuddy	V									V	
37.	Pham et al., 2018		V					English Practice AI chatbot was developed with Google's Dialogflow	V										V
38.	Setiawan et al.,2022		V					Interactive English Learning chatbot	V								V		
39.	Wang et al., 2017		V					a 3D virtual world, OpenSimulator with chatbot and time machine	V										V
40.	Shin et al., 2024		V					AI chatbot developed with Google's Dialogflow.	V								V		
41.	Wu et al., 2023		V					SpokenBot	V							V			
42.	Yang et al., 2022		V					Ellie, a voice AI chatbot	V								V	V	
43	Kohnke, 2022		V					AI chatbot developed with Google's Dialogflow.	V							V			
44.	Chiaráin & Chasaide, 2016		V					AI chatbot named Taidhgín developed using Pandorabots with ABAIR synthetic voices.						V			V		
45.	Bailey & Almusharraf, 2021			V				The AI chatbot integrated with Facebook Messenger was developed with the ActiveChat.ai chatbot platform.	V							V			
46.	Barley et al., 2021			V				AI storybot integrated with Facebook Messenger were developed with	V							V			

No.	Authors	Type of AI Chatbot						Platform / Application	Language							Subject				
		C	S	W	O	M	I		E	NM	FL				T	UG	MH	P	NM	
											G	M	J	OL						
								the ActiveChat.ai chatbot platform and Google's Dialogflow.												
47.	Chen et al., 2020			V				An AI chatbot named Xiaowen was developed using Python and the LINE Developer platform.				V				V				
48.	De Iacovo et al., 2021			V				AI chatbot integrated with Telegram was developed.						V		V				
49.	Haristiani et. al., 2022			V				An AI chatbot named Gengobot is integrated with LINE.					V				V			
50.	Haristiani & Rifa'I, 2020			V				An AI chatbot named Gengobot is integrated with LINE.					V			V				
51.	Haristiani & Rifa'I, 2021			V				An AI chatbot named Gengobot is integrated with LINE.					V			V				
52.	Johnson et al., 2022			V				Integrated with Telegram, an AI chatbot named Escapeling, is a sci-fi escape room.	V								V			
53.	Khlaisang & Sukavatee, 2023			V				The AI chatbot is incorporated into Facebook Messenger for reading and writing, whereas the iReview system is used for listening and speaking.	V							V				
54.	Kohnke, 2023			V				AI chatbot was developed with	V							V				

No.	Authors	Type of AI Chatbot						Platform / Application	Language						Subject				
		C	S	W	O	M	I		E	NM	FL				T	UG	MH	P	NM
											G	M	J	OL					
								Google's Dialogflow and integrated with Facebook Messenger.											
55.	Rodosthenous et al., 2020			V				LingoGameBot, an AI chatbot, integrates with Telegram. It is a part of the V-TREL system to offer vocabulary training exercises generated from the common-sense knowledge-based ConceptNet.	V							V			
56.	Oralbayeva et al., 2022			V				Reinforcement learning AI chatbot with integration of Telegram						V		V			
57.	Wu et al., 2024			V				Mandarin Second Language Intelligent Personal Assistant (MSLIPA), an intelligent personal assistant integrated with WeChat				V						V	
58.	Lee & Jeon, 2022			V				Voice-controlled conversational agents (VCAs) were developed using Google's Dialogflow and integrated into Google Assistant,	V										V	

No.	Authors	Type of AI Chatbot						Platform / Application	Language						Subject				
		C	S	W	O	M	I		E	NM	FL				T	UG	MH	P	NM
											G	M	J	OL					
								operating on a Google Home Mini device.											
59.	Alrajhi, 2024				V			A web-based AI pedagogical chatbot named Tutor Mike	V							V				
60.	Çakmak, 2022				V			AI chatbot Replika	V							V			
61.	Duong & Suppasetseree, 2024				V			Andy English Chatbot	V							V			
62.	Fathi et al., 2024				V			Andy English Chatbot	V										V
63.	Fryer et al., 2017				V			Cleverbot (Carpenter, n.d.) Chatbot	V							V			
64.	Fryer et al., 2019				V			Cleverbot (Carpenter, n.d.) Chatbot	V							V			
65.	Lin & Mubarok, 2021				V			AI chatbot Replika, with mindmap-guided flipped learning	V							V			
66.	Mu & Sarkar, 2019							SHRDLURN, an interactive language-learning game		V						V			
67.	Nozhovnik et al., 2023				V			A chatbot-driven e-classes on the Smart Sender platform	V							V				
68.	Annamalai, et. al., 2023					V		Several AI chatbots apps (Duolingo, Mondly & Andy).	V							V				
69.	Annamalai, et. al., 2023					V		Several AI chatbots apps (Duolingo, Mondly, Andy and Memrise).	V							V				

No.	Authors	Type of AI Chatbot						Platform / Application	Language						Subject				
		C	S	W	O	M	I		E	NM	FL				T	UG	MH	P	NM
											G	M	J	OL					
70.	Annamalai, et. al., 2023					V		Several AI chatbots apps (Duolingo, Mondly & Andy).	V							V			
71.	Belda-Medina & Calvo-Ferrer, 2022					V		Several AI chatbots apps (Replika, Kuki and Wysa)	V						V	V			
72.	Belda-Medina & Kokošková, 2023					V		Several AI chatbots apps (Mondly, Andy, John Bot, and Buddy.ai)	V						V	V			
73.	Coniam, 2008					V		Several AI chatbots apps (Cybelle, Dave, George, Jenny, Lucy)	V							V			
74.	El Shazly, 2021					V		Several AI web chatbots apps (Audrey, Charles, Cristal, and Mike) and Mondly, written and oral communicative virtual partners	V							V			
78.	Fryer & Carpenter, 2006					V		Chatbot Alice and Jabberwacky		V									V
79.	Kemelbekova, et. al., 2024					V		Text Cortex, a chatbot-driven AI system for vocabulary enrichment, Elsa Speak for voice recognition refinement and Grammarly for enhancing English grammar.	V						V	V			
80.	Kim et al., 2021					V		Several AI chatbots apps (Replika, Andy, and Google	V							V			

No.	Authors	Type of AI Chatbot						Platform / Application	Language							Subject				
		C	S	W	O	M	I		E	NM	FL				T	UG	MH	P	NM
											G	M	J	OL					
								Assistant)											
81.	Mohamed & Alian, 2023					V		Two AI chatbots (Duolingo and Falou)	V								V		
82.	Alm & Nkomo, 2022					V		Several AI chatbots apps (Duolingo, Memrise, Eggbu n and Mondly.)		V									V
83.	Chuah & Kabilan, 2021						V	Using AI chatbots Andy and Wordsworth, which are integrated with Telegram and Facebook Messenger	V						V				
84.	Ye et al., 2022						V	Microsoft Xiaoying, a chatbot launched by Microsoft on WeChat	V								V		
85.	Lorenzo et al., 2013					V		SLRoute, a combination of OpenSim with Alicebot and Pandora bots						V		V			
86.	Tai & Chen, 2020					V		Google Home Hub and its associated app, Google Assistant.	V								V		
87.	Zhang et al., 2024							AI chatbots		V						V			

C ChatGPT, S Self-developed AI chatbot without integration of social media applications/platforms, *W* Self-developed chatbot with the integration of social media applications/platforms, O Use only one AI chatbot available, *M* Use more than one AI chatbot available, *I* Integration of AI chatbot available with social media applications/platforms or virtual environment, *E* English, *FL* Foreign language, *G* German, *M* Mandarin/Chinese, J Japanese, *OL* Other languages, *UG* Undergraduate and above, *MH* Middle and High School, *P* Primary school and below, *NM* Not mentioned, *PEOU* Perceived Ease of Use, *PU* Perceived Usefulness, *BI to use* Behavioural Intention to Use, *AU* Actual Use, *PSQ* Perceived System Quality, *OCD* Online Course Design, *PE* Perceived Enjoyment, *PSE* Perceived Self-Efficacy, *SN* Subjective Norm, *EFL* English as a Foreign Language, *ESL* English as a Second Language.