ReCALL (2025), 37: 1, 4–21
doi:10.1017/S0958344024000168

RESEARCH ARTICLE

The use of artificially intelligent chatbots in English language learning: A systematic meta-synthesis study of articles published between 2010 and 2024

Fatma Şeyma Koç
Akdeniz University, Türkiye (fatmaseymakoc@akdeniz.edu.tr)

Perihan Savaş
Middle East Technical University, Türkiye (perihans@metu.edu.tr)

Abstract

In this qualitative systematic meta-synthesis study, 57 studies from the international literature published between 2010 and 2024 on the use of voice-based artificially intelligent chatbots in English language learning were analyzed. The present study aimed to explore the most recent studies on this topic by investigating the theoretical frameworks, methodological and technological properties, user reports of chatbot usage experience, and pedagogical implementations. It sought to identify research and implementation trends for voice-based chatbots via qualitative data analysis methods. Based on the reviewed studies, this paper presents data-based pedagogical implications that align with the latest voice-based AI chatbot research trends.

Keywords: English language learning; voice-based AI chatbots; artificially intelligent agents; improving language skills; meta-synthesis

1. Introduction

The proliferation of artificial intelligence (AI) in natural language processing has guided the construction of intelligent chatbots gifted at comprehending and generating human language (Caldarini, Jaf & McGarry, 2022). The popularity of devices like laptops, wearable tech, and smart home gadgets has expanded the use of readily accessible innovative technologies such as chatbots, embodied agents, and conversational agents like Alexa and Siri (Kukulska-Hulme & Lee, 2020). In their earlier versions, chatbots communicated with users in a textual mode through keyword matching (Jia, 2009). However, in the 1980s, chatbots with speech recognition and text-to-speech capabilities were developed (Godwin-Jones, 2023). The combination of automatic speech recognition and text-to-speech technologies in AI-powered chatbots has resulted in a mechanism capable of conversing naturally and interactively with learners (Belda-Medina & Kokošková, 2023). Although both text- and voice-based modalities have been employed towards learning goals, each has had different use venues. Voice modality chatbots have been mainly used in language classes to transform classroom experience to construct aural input and opportunities for oral interaction. Learners are increasingly adopting voice-based chatbots because, compared to

Cite this article: Koç, F.Ş. & Savaş, P. (2025). The use of artificially intelligent chatbots in English language learning: A systematic meta-synthesis study of articles published between 2010 and 2024. *ReCALL* 37(1): 4–21. https://doi.org/10.1017/S0958344024000168

© The Author(s), 2024. Published by Cambridge University Press on behalf of EUROCALL, the European Association for Computer-Assisted Language Learning. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

https://doi.org/10.1017/S0958344024000168 Published online by Cambridge University Press

text-based interaction, the voice-based interaction modality is more akin to face-to-face interaction and has certain benefits over text-based communication. Namely, it exposes learners to paralinguistic elements like stress, intonation, and other suprasegmental components (Rassaei, 2023).

Thus, the adoption of voice-based AI chatbots in English language learning has surged in recent years (Hwang, Guo, Hoang, Chang & Wu, 2022; K.-A. Lee & Lim, 2023; S. Lee & Jeon, 2022; Park, 2022; Timpe-Laughlin, Sydorenko & Daurio, 2022; Wu, Lam, Kong & Wong, 2023) and their use is becoming more widespread in line with the extensive use of portable computer-assisted language learning (CALL) and mobile-assisted language learning (MALL) devices (Kukulska-Hulme & Lee, 2020). However, more recent and comprehensive evidence is still needed on implementing voice-based AI chatbots in language classes, as new voice-based and AI-supported chatbot tools are being introduced almost weekly with little or no prior research or piloting in educational contexts. Educators especially are in need of continuous input, support, and guidance in order to keep up with this dynamic field and the use of chatbots, which can be achieved via dynamic and systematic research. In particular, demonstrating how voice-based AI chatbots are used in various countries at different levels of institutions can provide English language teachers and future researchers with guidelines and insights about the practical characteristics of AI chatbots for particular groups of learners. There is also a need for more guidance for English language teachers on the task, material, and activity design for integrating voice-based AI chatbots into their classes. In addition, even though voice-based AI chatbots are gaining popularity with a general audience, there are still areas that need to be improved in terms of context and culture-specific discourse formation; therefore, more research in this area can give insight into the development of pedagogically sound and data-driven voice-based chatbots geared primarily for language classes. This study aims to demonstrate a clear picture of the theoretical grounds of voice-based AI chatbot research conducted in many countries worldwide and illuminate learning and affective outcomes in line with specific characteristics of AI chatbots employed in various educational institutions.

Thus, a meta-synthesis of the studies mentioned above that center around voice-based AI chatbots might provide insights into recognizing what these technologies offer for English language learning at both micro and macro levels by analyzing each study in detail and also by providing a holistic picture presenting the synthesis of the studies carried out in recent years. Up to this date, several review studies on AI chatbots have been implemented (Huang, Hew & Fryer, 2022; Ji, Han & Ko, 2023; Zhai & Wibowo, 2022), but these review studies have different focal points. Huang *et al.* (2022) analyzed 25 studies from 2008 to 2020 based on chatbots' technological, pedagogical, and social affordances for language learning in first-language and foreign-language learning contexts. Zhai and Wibowo (2022) explored the empathy, human, and cultural dimensions of using chatbots in language learning. Moreover, Ji *et al.*'s (2023) review of 24 studies focused on the role of collaboration with teachers in AI chatbot–supported language classes. In addition, two systematic reviews have recently been published, particularly on voiced-based AI chatbots in language learning. Jeon, Lee and Choe (2023) analyzed the studies based on goal orientation, embodiment, and multimodality in their review. Jeon, Lee and Choi (2023) explored general research trends and the role of voice-based chatbots (i.e. feedback providers) as language learning resources. The present meta-synthesis aims to further the investigation by exploring other dimensions and synthesizing the findings of the reviewed studies on voiced-based chatbots in terms of linguistic and affective factors, theoretical frameworks, technological implementations, and pedagogical implications. In line with the aims of this study, the following research questions were generated.

1. What theoretical backgrounds are employed in AI chatbot–integrated studies in the English language learning domain published between 2010 and 2024?
2. How can the literature on voice-based AI chatbot research be characterized regarding methodology between 2010 and 2024?

https://doi.org/10.1017/S0958344024000168 Published online by Cambridge University Press

3. What technologies and/or chatbots are used in the published research between 2010 and 2024?
4. What strengths and challenges were reported in the studies published between 2010 and 2024?
5. What pedagogical implementations were prevalent in the studies published between 2010 and 2024?

2. Methodology

The meta-synthesis presented in this paper employed the PRISMA framework (Preferred Reporting Items for Systematic Reviews and Meta-Analyses), a protocol by Moher, Liberati, Tetzlaff, Altman and the PRISMA Group (2009) to provide authors with explicit criteria for transparent and effective reporting. The methodology of this systematic meta-synthesis involves three main steps: searching for articles, selecting articles, and qualitative coding of data from selected articles.

2.1 Search strategy

Following the guidelines of the PRISMA framework to carry out a qualitative systematic review study, a rigorous search strategy was adopted. The key terms "chatbot," "conversational agent," "conversational system," "dialogue-based," "conversational dialog agent," "human-computer dialog," "language learning," "English learning," "foreign language," "second language," "L2," "EFL (English as a Foreign Language)," "ESL (English as a Second Language)," "language acquisition," "CALL," and "MALL" were used with a mixture of Boolean expressions "AND" and "OR." While in some studies researchers created web-based AI chatbots (Ayedoun, Hayashi & Seta, 2020; El Shazly, 2021), others developed app-integrated AI chatbots (Belda-Medina & Calvo-Ferrer, 2022). Therefore, since researchers have used a variety of concepts to describe the utilization of AI chatbots depending on their digital format, we decided to include both "CALL" and "MALL" as keywords in our article search of databases. All keywords were searched in journal names, abstracts, and full articles in databases ranging from Web of Science, ERIC, Scopus, Elsevier, EBSCOhost, and Springer to access articles, book chapters, and conference proceedings to be reviewed matching our criteria for selection. In addition to the articles retrieved from these databases, articles found manually in the reference sections of the articles were also included. During the writing and revision process of the paper presented here, some of the initially reviewed studies (Chen, Yang & Lai, 2023; M.-H. Hsu, Chen & Yu, 2023; Jeon, 2024) that were published with a DOI number and with early access date were published in the journals with an issue number at a later date. The present study's authors updated the volume, issue numbers, and publishing dates. The selected studies for review have been shown with an asterisk in the reference section. The whole list of selected studies for review is given in the supplementary materials.

2.2 Selection criteria for the studies

Conceptual, theoretical papers, articles using languages other than English, studies outside of the English language teaching field, and studies that do not report the experiences of chatbot users were eliminated. The works on voice-based AI chatbots and English language learning were the main focus of the search. Under AI chatbots, we also included studies using intelligent personal assistants (IPAs) such as Google Assistant and Alexa because IPAs are also branched under closed-domain retrieval-based AI chatbots (H. Kim, Yang, Shin & Lee, 2022).

https://doi.org/10.1017/S0958344024000168 Published online by Cambridge University Press

Although this meta-synthesis study is based on the English language learning experiences of chatbot users, feasibility studies using state-of-art technologies in creating this much-needed technology have also been included, as one dimension in the study was to explore these tools' technical qualities.

This study's focus on the time frame from 2010 to 2024 derives from our intention to concentrate on studies utilizing voice recognition chatbots with the latest AI, automatic speech recognition (ASR), and natural language processing technologies. The literature review showed that although the development of voice-based AI chatbots dates back to the 1980s, the research on voice-based chatbots in students' English language learning experience did not surge until after 2016, as we could trace back to only one study published each year during 2010 and 2012. Table 1 demonstrates the number of papers reviewed and analyzed by journal name and year of publication (see supplementary materials).

2.3 Data extraction and analysis of the studies selected

Using PRISMA guidelines, an article search was carried out, and 57 articles were selected based on our selection criteria. Figure 1 outlines the selection of studies following PRISMA guidelines (Moher *et al.*, 2009). The constructivist grounded theory (CGT) by Thornberg and Charmaz (2014) was used to analyze the findings and research trends in the studies. The main results of the AI chatbot studies were initially extracted in the form of raw data. The data were first read line by line in the initial coding phase and coded through a constant comparison method to examine data in contrast to other data, analyze data in contrast to codes, and compare code with code to identify any arising discrepancies or congruities. Subsequently, the major findings in the reviewed studies were specified and synthesized through a qualitative coding process while interpreting the emergent codes in the focused coding stage.

Considering every article as an analytic component, we comprehensively and iteratively analyzed all the articles. First, we read the literature on reviews of chatbots in general (e.g. Pérez, Daradoumis & Puig, 2020) to create a scheme for a deductive coding scheme. We constructed an initial coding framework of the methodology, settings, and participants based on the previous reviews. Through data-driven inductive coding, four further categories were constructed: theoretical frameworks, strengths and challenges, technologies, and pedagogical implications. These dimensions were deemed particularly important to this study, as examining these factors can provide insights into the technical-pedagogical aspects of voice-based AI chatbots used in past studies. The findings can be employed to understand the implications of using such chatbots with particular student profiles in various educational settings.

This study followed the CGT in the data analysis, as this theory allowed the enactment of an interpretative research hypothesis with a cyclical data analysis process through constant data comparison. This process enabled us to co-construct meaning during the analysis phase instead of validating a priori hypothesis. The selected studies for the present study were systematically synthesized through a meta-synthesis. The rationale behind using a meta-synthesis approach was that we aimed to reach a novel and unifying understanding of findings that was more substantial and profound than a single study and to get a broader understanding of the use of voice-based AI chatbots in English language learning (Finfgeld, 2003; Saini & Shlonsky, 2012). These unified findings were illustrated as an analysis table in Table 2 (see supplementary materials). After reaching a consensus on the data extraction form by two researchers, one researcher extracted related information from the reviewed studies. The validity of this process was ensured via data extraction of the reviewed studies by one other researcher.

https://doi.org/10.1017/S0958344024000168 Published online by Cambridge University Press

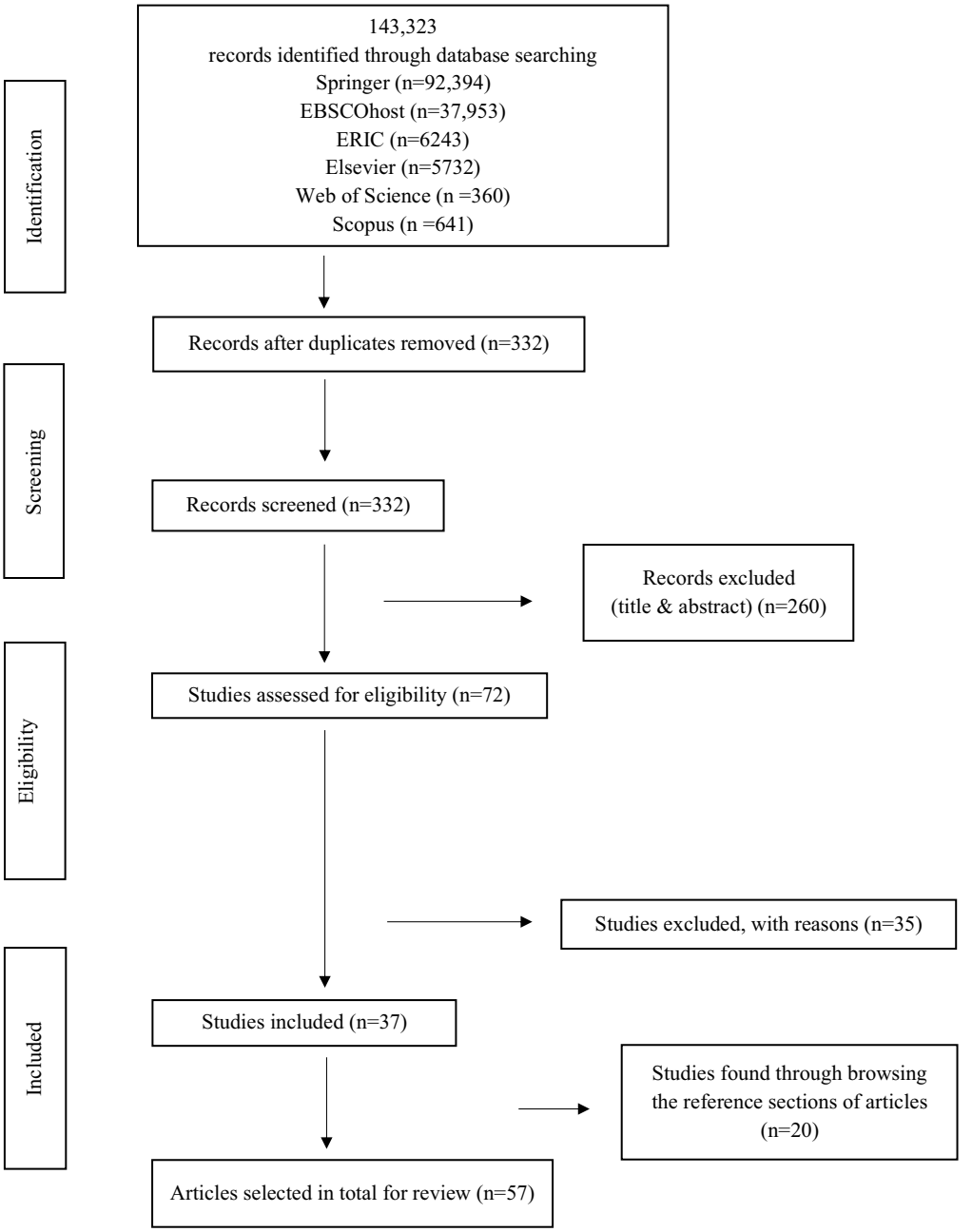


Figure 1. PRISMA flow chart (Moher *et al*., 2009).

3. Findings

3.1 Theoretical backgrounds in the reviewed studies

The first category is theoretical backgrounds, which comprise the theories employed in chatbot-assisted language learning research (see Table 3 in supplementary materials). Namely, theoretical frameworks point to theoretical foundations of AI chatbot research targeting English language learners at K–12 and higher education levels, which comprise linguistic, cognitive, social, and

https://doi.org/10.1017/S0958344024000168 Published online by Cambridge University Press

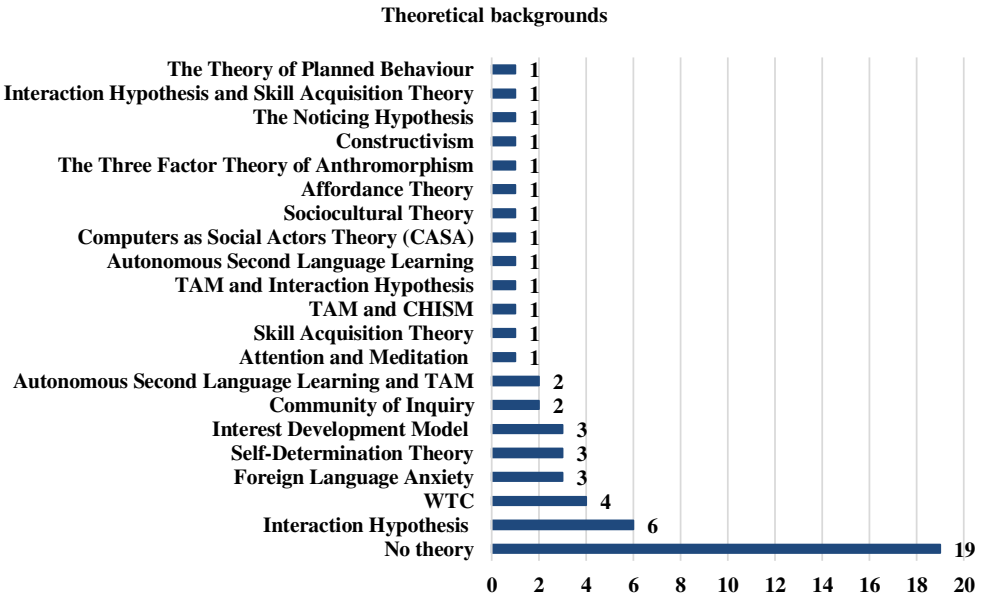


Figure 2. Theoretical backgrounds.
Note. TAM = technology acceptance model; CHISM = chatbot–human interaction satisfaction model; WTC = willingness to communicate.

technological-pedagogical theories. The linguistic category is related to language learning theories (i.e. the input hypothesis). The cognitive category comprises cognitive learning theories like the interest development model. The social category refers to social learning theories such as the constructivist theory. At the same time, other theories explaining general technological frameworks applied to studies, such as community of inquiry (CoI), were categorized as technological-pedagogical.

As displayed in Figure 2, linguistic frameworks were the most frequently adopted theoretical backgrounds ($n = 21$), followed by technological-pedagogical ($n = 11$), cognitive ($n = 4$), and social ($n = 2$) theories. More particularly, the interaction hypothesis ($n = 8$), willingness to communicate (WTC) ($n = 4$), autonomous second language learning ($n = 3$), foreign language anxiety (FLA) ($n = 3$), and skill acquisition theory ($n = 3$) were most frequently used for grounding theories in studies.

Based on Long's (1996) interaction hypothesis, studies sought to discover if chatbots would increase interaction in L2 (i.e. referring to a second language) by offering favorable circumstances for the negotiation of meaning, interaction, input, and output (Dizon, 2017, 2020). The studies reviewed were also based on the claim that the use of technology might help increase the WTC levels of learners – that is "a readiness to enter into discourse at a particular time with a specific person or persons, using an L2" (MacIntyre, Clément, Dörnyei & Noels, 1998: 547), and it has been found that learners demonstrated a higher expected WTC after talking to a chatbot (Ayedoun, Hayashi & Seta, 2019a; Ayedoun *et al.*, 2020; Tai & Chen, 2023). Other studies also based their arguments on Gardner and Miller's (1999) self-access language learning theoretical framework, claiming that learners' autonomy and independence can be promoted through the use of chatbots as supplementary language learning materials (Dizon & Tang, 2020; Moussalli & Cardoso, 2020).

Another salient theory used by researchers was the interest development model (Hidi & Renninger, 2006). Following this theory, researchers investigated whether AI chatbots could trigger students' interest in language learning (Fryer, Ainley, Thompson, Gibson & Sherlock, 2017;

https://doi.org/10.1017/S0958344024000168 Published online by Cambridge University Press

Fryer, Nakao & Thompson, 2019; Fryer, Thompson, Nakao, Howarth & Gallacher, 2020). Basing their studies on the self-determination theory, researchers also explored if chatbots would facilitate autonomy, competence, and relatedness feelings of language learners (Annamalai *et al.*, 2023; Jeon, 2022; Morton & Jack, 2010; Ryan & Deci, 2017). Moreover, based on the theory of FLA, researchers have examined the impact of using chatbots on learning anxiety (Dizon, 2017; Horwitz, Horwitz & Cope, 1986; Tai & Chen, 2023) and the meditation and relaxation levels of language learners (L. Hsu, 2022).

The investigation of chatbot usage has also been carried out under the principles of the skill acquisition theory (Lyster & Sato, 2013), revealing that chatbots would be constructive tools for language learners thanks to their potential to provide structured oral practice (Sydorenko, Smits, Evanini & Ramanarayanan, 2019). Studies also sought to discover the potential of AI-powered chatbots for fostering oral interaction by raising awareness of specific L2 pragmatic structures based on Schmidt's (1995) noticing hypothesis (Timpe-Laughlin & Dombi, 2020). Moreover, Ericsson, Sofkova Hashemi and Lundin (2023) also grounded their studies on the sociocultural theory, arguing that interaction with AI chatbots might provide instances for social interaction through human–computer dialogue (Lantolf, Thorne & Poehner, 2014). Based on the constructivist theory in technological learning settings, Park (2022) explored and identified the role of a chatbot in promoting student-centered and scaffolded learning as a learning tool for geographically remote learners (Jonassen, Davidson, Collins, Campbell & Haag, 1995).

Besides the linguistic and social theories, learners' perspectives in using chatbots for language learning concerning factors such as user engagement and linguistic and technological components were explored through the technology acceptance model (TAM) and the chatbot–human interaction satisfaction model (CHISM) (Belda-Medina & Calvo-Ferrer, 2022; Davis, 1989). A similar theory, the theory of planned behavior by Fishbein and Ajzen (1975), was adopted by Wu *et al.* (2023) to analyze the relationship between the factors of the behavioral intention of using chatbot applications and learner beliefs, such as the usefulness of voice-based AI chatbots for learning.

The computers as social actors (CASA) theory by Nass, Moon, Fogg, Reeves and Dryer (1995) has also been adopted to explore how AI chatbots affect language learners' learning motivation. Researchers aimed to observe whether AI chatbots would create a feeling of social presence among learners in cases where they are described as having human-like properties (Ebadi & Amini, 2024). Epley, Waytz and Cacioppo's (2007) three-factor theory of anthropomorphism, consisting of "elicited agent knowledge, effectance, and sociality" factors, referring to knowledge about agents, motivation to interact with them, and intention to communicate with agents, respectively, was also used to analyze if AI chatbots were described with human-like characteristics (S. Lee & Jeon, 2022: 3). Adopting the affordance theory by Gibson (1986), researchers also analyzed the experiences of language learners in using voice-based AI chatbots in terms of the technological, pedagogical, and social contributions to language learning (Jeon, 2024). Furthermore, the CoI framework by Garrison and Arbaugh (2007) has also been used in studies to observe the effect of the social, cognitive, and teaching presence on social dynamics, learning outcomes, and learning performance of language learners after interacting with voice-based AI chatbots (Wang *et al.*, 2023; Wang, Pang, Wallace, Wang & Chen, 2022). However, 19 studies were not based on any theoretical groundings. Therefore, a clear-cut theoretical framework still needed to be included in many studies.

3.2 Methodological trends in the reviewed studies

The second category is the methodologies, consisting of four variables: design features, countries, education level, and sample size. Regarding the methodology, there were 4,062 participants in these 57 studies. In 23 of the studies using an experimental methodology, data were collected from 1,408 participants. Two studies used sequential explanatory mixed methods with 432 participants.

https://doi.org/10.1017/S0958344024000168 Published online by Cambridge University Press

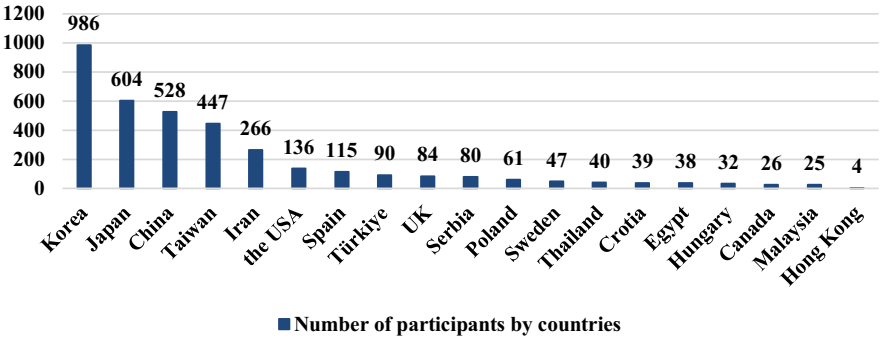


Figure 3. Research regions.

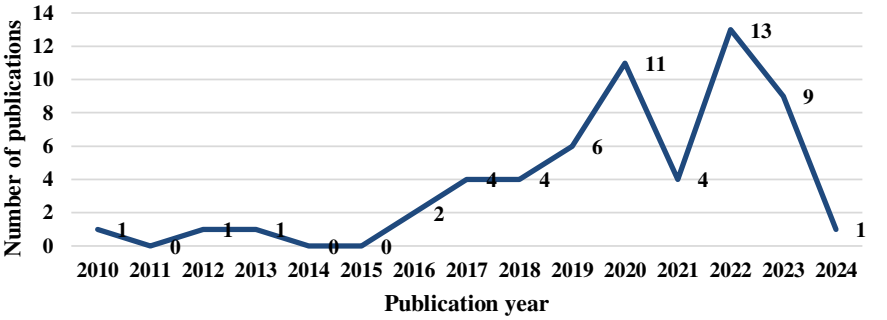


Figure 4. The course of voice-based AI chatbot publications over the years.

Ten feasibility studies with 246 participants were also conducted, and the five case studies comprised 57 participants. However, the remaining studies did not specify their methodologies explicitly, and they had 1,674 participants. Most participants were undergraduate and graduate students (49%, $n = 1{,}984$), while high school students constituted 23% of the population ($n = 947$), and primary school students constituted 23% of the population ($n = 951$). The remaining studies did not specify the education level of the students.

The meta-synthesis study presented here included studies conducted in various countries. However, most studies, namely 13 out of 57, have been conducted in Japan. The sample consisted of a total of 986 participants from Korea, followed by Japan ($n = 604$), China ($n = 528$), and Taiwan ($n = 447$), respectively. Figure 3 below displays the distribution of the participants in the reviewed studies by countries.

Moreover, the publication dates of the reviewed articles were evenly distributed over six years, with an upward trend in the number of publications beginning in 2016. Figure 4 denotes the trend for voice-based AI chatbot research in English language teaching over the years.

3.3 The types of technologies and/or chatbots used in the studies analyzed

The next category of analysis is the types of voice-based AI technologies, which comprised three codes: general audience chatbots (i.e. Alexa, Siri, Cortana, etc.), specific-purpose chatbots created for language learners, and AI chatbots embedded in virtual learning environments. Twenty-seven types of chatbots were identified (see Figure 5).

The results showed that the these chatbots had been adopted to support English language learning. The most common practice among the synthesized studies was employing already

https://doi.org/10.1017/S0958344024000168 Published online by Cambridge University Press

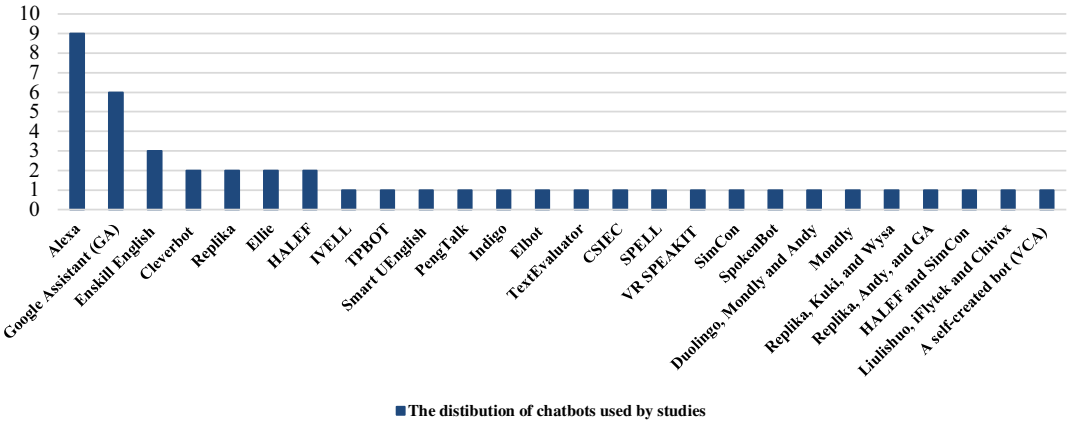


Figure 5. The distribution of chatbots used by studies.

present open-access or commercial applications created for native speakers or L2 learners. Thus, two such chatbots, Alexa and Google Assistant, were used the most. Two studies among the reviewed studies also employed Cleverbot, an open-access chatbot (Fryer *et al.*, 2017, 2019), which is an updated version of Jabberwacky, a Loebner prize winner for producing the most similar output to human speech.

Another trend in the synthesized studies was the creation of unique AI chatbots by the authors and testing their effectiveness on language learners. However, these chatbots were created for academic purposes and were not publicly available. For instance, Ayedoun *et al.* (2019a, 2019b, 2020) designed DiMaCA, a dialogue management model consisting of the dialogue flow management and the strategies management modules to create a chatbot, Peter, which used affective backchannels and communication strategies. Sydorenko *et al.* (2019) also created a voice-based HALEF application targeting students of English as a second language.

The results also showed that chatbots aiming at lower language proficiency groups were rare in the studies reviewed. In one study, TextEvaluator was created targeting third- and fifth-grade English language learners in the USA (Forsyth *et al.*, 2019); another chatbot, SPELL, was created for junior high school students in China (Morton & Jack, 2010), and one other chatbot was designed for elementary school learners in China (Wang *et al.*, 2023). Additionally, in a few studies, chatbots were embedded in virtual learning environments (Hassani, Nahvi & Ahmadi, 2016; Morton, Gunson & Jack, 2012; Park, 2022), yet the integration of virtual reality technology into AI chatbots was relatively uncommon among the reviewed studies.

3.4 Strengths and challenges in the use of chatbots in the reviewed studies

Lastly, strengths and challenges were determined by analyzing the results and discussion parts of the reviewed articles. The strengths and challenges discovered in the reviewed studies are illustrated in Table 4 (see supplementary materials). The most commonly encountered strengths in studies were developing language skills ($n = 27$), enhancing affective factors ($n = 27$), facilitating communicative competence ($n = 14$), and increasing pragmatic awareness and competence ($n = 10$). It has been discovered in our synthesis that language learners can develop their linguistic skills to a great extent thanks to the interaction with voice-based AI chatbots during English language learning. Significant gains in speaking scores (Dizon, 2020; El Shazly, 2021; D.-E. Han, 2020; M.-H. Hsu *et al.*, 2023; Hwang *et al.*, 2022; Moussalli & Cardoso, 2021; Park, 2022; C. T.-Y. Yang, Lai & Chen, 2022) and listening scores (N.-Y. Kim, 2018; Tai & Chen, 2022) were reported after engaging in negotiating meaning with chatbots.

https://doi.org/10.1017/S0958344024000168 Published online by Cambridge University Press

The findings of the analyzed studies have also demonstrated positive aspects for affective factors in the use of chatbots by English language learners: higher attention and meditation levels, reduced anxiety, higher WTC, and motivation levels after interacting with voice-based AI chatbots (K.-A. Lee & Lim, 2023; Tai & Chen, 2023; C. T.-Y. Yang *et al.*, 2022). For instance, L. Hsu (2022) carried out a neuroscientific experimental study using a chatbot they created and tested it on EFL learners to analyze their attention – the level of mental "focus" and meditation and calmness/relaxation states – by measuring their brainwave activities when they practice speaking in three different environments (face-to-face, virtual human to human, and human to an AI chatbot conditions) to explore the interlocutor effect. Meditation level was the highest in human chatbot conditions, which implied the beneficial effects of chatbots in reducing English language learning anxiety.

The synthesized studies also report the challenges of using chatbot systems. The low intelligibility of learner utterances due to problems with speech recognition ($n = 15$), the unnaturalness of AI chatbot–human interaction ($n = 9$), and a lack of explicit corrective feedback ($n = 5$) were the most commonly reported problems. These challenges were faced due to various issues. According to the studies analyzed, learners appeared less motivated to converse with chatbots than human partners due to novelty factors and unnaturalistic speaking conditions (Ericsson, Lundin & Sofkova Hashemi, 2023; Fryer *et al.*, 2017). For instance, the results of a study by Belda-Medina and Calvo-Ferrer (2022) showed only moderate interest in using AI chatbots for a more extended period. Furthermore, Çakmak (2022) and El Shazly (2021) identified increased levels of anxiety after interacting with the AI chatbots, which was attributed to factors such as the educational context that could ignite worries about failing the course as well as self-consciousness or language ego of students exposed to error correction, and lack of empathy and complex reasoning in chatbots.

3.5 Pedagogical implementations in the studies analyzed

The last category pertained to the pedagogical implementations, which are concerned with guiding and helping learners use AI chatbots to interact as conversational partners in learning languages through activities such as training and scaffolding and designing conversational scenarios that best fit this purpose. Chatbots have substantial potential for teaching languages, yet there is a need for more guidance on classroom implementations. Therefore, the researchers' use of these tools in classroom settings is relevant to this study. Learners need guidance to utilize chatbots effectively for language learning purposes. An analysis of the reviewed studies on pedagogical practices in AI chatbot implementation has shown that two strategies guide AI chatbot users: (1) providing structured or semi-structured prompts, and (2) providing linguistic strategies.

First, regarding the prompts, researchers have discovered that several strategies were implemented to direct learners to utilize AI chatbots for English language learning. The analyzed studies displayed that teachers applied methods containing structured and semi-structured prompts for learners' guidance during the conversations. Students were given topics, questions, and hints in question-and-reply stems (Ayedoun *et al.*, 2019a, 2020; Fryer *et al.*, 2017, 2019, 2020; L. Hsu, 2022; Johnson, 2019; Tai & Chen, 2023). Furthermore, while interacting with chatbots, learners were given prompts of pre-established questions and language commands such as "Play me a song/joke/story" and instructed to ask chatbots like Google Home Hub and Alexa other general-knowledge-related facts such as the weather, temperature, and the capital of the countries. This way, students could learn to utter appropriate commands during conversations (Chen *et al.*, 2023; Dizon, 2017; Moussalli & Cardoso, 2020).

The synthesized studies also reported that students were provided with linguistic strategies in the pre-task phase before using the chatbots to familiarize them with these tools. The strategies they were suggested to use to get into a dialogical conversation included reflecting on the misunderstood aspects of their speech, modifying their output by using more precise and slower speech, altering the pronunciation of words (Dizon, 2020), and asking for the definition, spelling,

https://doi.org/10.1017/S0958344024000168 Published online by Cambridge University Press

synonym, or the translation of the words to be able to notice the linguistic gaps in their knowledge in the face of a communication breakdown (Chen *et al.*, 2023; Dizon, 2017; Moussalli & Cardoso, 2020).

Another theme regarding the pedagogy of AI chatbots in English language learning was the design methodology of pedagogical activities in the reviewed studies. Our review illustrated that researchers used the principles of Ellis's (2003) task-based language teaching using tasks in speaking scenarios for language practice in designing AI chatbots (Ayedoun *et al.*, 2019a, 2020; Li, Chang & Wu, 2020; Sydorenko, Daurio & Thorne, 2018; Sydorenko *et al.*, 2019). To illustrate, in a study using HALEF, the users practiced speaking English in a coffee shop simulation. The chatbot and the users interacted to complete an order (Sydorenko *et al.*, 2018, 2019; Timpe-Laughlin & Dombi, 2020). Similarly, students talked to an animated character in task-based simulation modules in Enskill English. The tasks included asking for directions and buying train tickets (Ericsson, Lundin & Sofkova Hashemi, 2023; Ericsson, Sofkova Hashemi & Lundin, 2023; Johnson, 2019). Thus, through partially scripted task-based or non-task-based dialogue activities in studies (Ayedoun *et al.*, 2019a, 2020; Johnson, 2019; Sydorenko *et al.*, 2019), students got engaged in two-way reciprocal communication with voice-based AI chatbots.

4. Discussion and implications

This study analyzed studies on integrating voice-based AI chatbots in the English language learning field. It aimed to demonstrate past research trends and suggest directions for future research on voice-based AI chatbots in English language learning.

4.1 An overview of the theoretical foundations of AI chatbot–integrated research in EFL/ESL contexts

The first research question concerned the theoretical focuses of the studies reviewed. In this study, most language learning activities through voice-based AI chatbots were guided by linguistic frameworks, as researchers might aim to construct their studies on these frames to clarify the interchange between characteristics of chatbot-supported language learning activities and learning outcomes.

Our findings demonstrated that several studies in AI chatbot–integrated language learning research had a basis in the interaction hypothesis (Long, 1996). This finding is unsurprising as there is insufficient input, especially in the EFL settings, and chatbots provide opportunities to be exposed to input in a foreign language (H. Yang, Kim, Lee & Shin, 2022). This study also provided evidence for other theories related to language development in AI chatbot–integrated language research. Namely, WTC, interest development, and FLA theories have been discovered to help explore chatbot-integrated language learning. This study demonstrated that in human chatbot conditions, learners had comparatively less anxiety about speaking in L2 (El Shazly, 2021; Tai & Chen, 2023). This finding implies that teachers can employ AI chatbots to aid learners in overcoming their anxiety about speaking in a foreign language. Among the reviewed studies, the higher WTC and interest levels due to interaction with chatbots suggested these tools' potential as motivators and sources of interest for language learners. Furthermore, this study demonstrated that self-regulated and autonomous learning theories could be influential in assisting and guiding chatbot-integrated language learning. According to self-regulated learning theories, when learners can regulate the content and way of learning and actively participate in learning activities, they can benefit most from the learning process (Reinders & White, 2016). Similarly, AI chatbots are tools that have the potential to promote language learner autonomy by providing ample opportunities for language practice without the need for the presence of a real speaking partner for language learners.

https://doi.org/10.1017/S0958344024000168 Published online by Cambridge University Press

Social learning theories have also been applied as the theoretical frames in voice-based AI chatbot research. Thus, based on the findings, it could be suggested that among the many reasons for adopting chatbot technology, one particular aim can be to create dialogical spaces for interaction using chatbots as mediating tools for language learning. Teaching professionals can also adopt numerous technological-pedagogical theories, such as CASA and CoI frameworks, to plan the language teaching process. The findings of this study imply that voice-based AI chatbots can serve as human-resembling language partners, as learners tend to attribute anthropomorphic characteristics to these learning tools. Language learners are most likely to benefit from using AI chatbots, as they positively affect learning engagement, mastery of the language, and constructive learning experiences.

All in all, various theories have been put forward to demonstrate the role of AI chatbots in L2 learning. On the other hand, only some studies referred to a solid theoretical background on which they conducted AI chatbot studies. Namely, 19 studies still needed to build their studies on theoretical backgrounds. This might indicate these researchers' tendency to center on other aspects of chatbot-supported language learning, such as intervention impacts (Hwang *et al*., 2022; N.-Y. Kim, 2018) and chatbot design features (Chen *et al*., 2023; K.-A. Lee & Lim, 2023). Without concrete theoretical backgrounds, it could be hard to determine which variables impact AI chatbot–supported language learning (AI CSLL) and explore how they can be incorporated into learning activities or curricula. Therefore, studies are suggested to refer to particular theories in future voice-based AI chatbot research. On the other hand, the extensive range of frameworks identified in the reviewed studies might also show that scholars are still hypothesizing how chatbot-supported learning and teaching could be theorized, suggesting that voice-based chatbot-supported language learning is at an earlier phase in its progress.

4.2 The necessity for more varied methodologies and contexts to carry out studies on speech-recognition chatbots

The second research question concerned research methodologies, settings, and participants. Regarding the educational level, we discovered that most research on AI chatbot use for English language learning was conducted in higher education settings. This result could be attributed to increased online and computer-assisted learning in higher education environments. This could also result from the accessibility of undergraduate students and the convenience of conducting chatbot studies with this group of learners. Undergraduate students might demonstrate more autonomous learning behaviors, as using chatbots for language learning might require self-regulated capabilities to cope with the requirements of this novel technology in and out of class. There were also fewer primary and junior high school participants. This result pointed towards a need for further studies to be carried out with groups of language learners, especially at primary levels of education, to explore the needs of an underrepresented population.

Furthermore, it was also discovered that a mixed-methods methodology was mainly adopted to represent stakeholders' opinions in the English language field and to demonstrate the learning outcomes of using voice-based AI chatbots. The prevalent use of quantitative and mixed methods in this study could be because of the impetus to comprehensively analyze the effects of AI chatbot technology. Studies with more varied methods are thus suggested to investigate the impacts of voice-based AI chatbots in English language learning.

When the geographical distribution of the countries was screened, it was seen that the studies were primarily implemented in the Asian context. The adoption of voice-based AI chatbots to support English language learning has gained more attention from researchers in certain Asian countries than elsewhere. This could be attributed to the fast incorporation of technology into language classes in these regions. Studies that cover other geographical areas might yield different results. Teachers and researchers in different parts of the world with different educational trends could conduct research in this field to provide comparable data.

https://doi.org/10.1017/S0958344024000168 Published online by Cambridge University Press

4.3 The ever-changing background of voice-based AI chatbot research

The third and fourth research questions were related to the technologies and chatbots adopted and the strengths and challenges experienced by English language learners. The findings demonstrated that researchers predominantly used open-access general audience chatbots such as Alexa and Siri by target language learner groups. Instead, there were fewer cases of creating voice-based AI chatbots specifically for English language teaching purposes. This result could point towards a need for more guidance for language teaching professionals to design specific-purpose AI chatbots.

The studies also addressed various strengths and challenges of using voice-based AI chatbots for English language learning. Our meta-synthesis shares similar findings to that of N.-Y. Kim, Cha and Kim's (2019) review in that studies demonstrated the effectiveness of AI chatbots in facilitating conversation, boosting interaction and negotiation of meaning, and elevating communicative competence, motivation, and engagement in English language learning tasks. On the other hand, the results of this meta-synthesis study align with Huang *et al.*'s (2022) and N.-Y. Kim *et al.*'s (2019) reviews, which report the narrow scope of the knowledge database, the unnatural voice of AI chatbots and chatbots' inability to comprehend users' language input as prevalent limitations. Some of the reasons behind these technical problems might be that although there is progress towards more functioning ASR and text-to-speech technologies, specific issues, such as inaccurate speech recognition and incomprehensibility of emotions (i.e. rhythm, intonation, and pitch), still need to be resolved (Belda-Medina & Kokošková, 2023).

These findings suggested a need to create more developed voice-based AI chatbots to address the various language learner profiles. Within this regard, one suggestion for application developers can be to train chatbots based on non-native speaker data, as most of the current chatbots are trained on native speaker corpus, which restricts the intelligibility of non-native utterances by the ASR system (Chen *et al.*, 2023). Our findings point out a need to improve the interactional capabilities of the present voice-based AI chatbots. Another suggestion is to develop their speech synthesis rather than relying on the present built-in synthesis on operating systems on mobile devices. Adding a distinctive voice that matches the personality and appearance of a chatbot as well as gestures and expressions are other suggested technical features expected to be inherent in chatbots (Yuan, 2023).

Based on the strengths and challenges reported in the AI chatbot–assisted language learning studies, another factor to be developed is error correction by the AI chatbot. Indisputably, personalized learning is one crucial goal that AI chatbot developers must seek. Therefore, one aim of future AI chatbot technologies should be to construct error correction and immediate feedback algorithm systems that can add to the personalization of the learning experience (Yuan, 2023). Therefore, AI chatbot creation is still an area that needs development.

On the other hand, even though chatbots still need many features to be added before being fully compatible with the aims of language learners, the increasing number of studies in the last few years can be attributed to the recent attention in the field and the particular benefits they provide, suggesting a further trend of interest in the following years.

4.4 Pedagogical implications in AI chatbot–integrated research

This study also specified perspectives regarding pedagogical implementations in voice-based AI chatbot research. The findings pointed towards the positive impact of the provision of strategies by educators prior to the use of the chatbot by language learners. The scaffolded learning support principle is one approach that could be implemented to help language learners participate in tasks conducted with voice-based AI chatbots. This principle consists of demonstrating and modeling communication with chatbots and providing cues to language learners to prevent interaction disruptions through meaningful negotiation strategies (J. Han & Lee, 2024). The provision of

https://doi.org/10.1017/S0958344024000168 Published online by Cambridge University Press

structured or semi-structured prompts or language commands as pre-task activities could equip learners with prior knowledge and strategies to be fully immersed in chatbot-assisted activities.

Designing the learning activities through task-based learning scenarios was another common and effective practice among the studies. This finding suggests that AI chatbots can be incorporated into language tasks through activities to provide language learners with authentic communication and instant feedback, contributing to L2 enhancement (Ellis, 2003; Xiao, Zhao, Sha, Yang & Warschauer, 2023). Employing voice-based AI chatbots based on a task-based design of the educational activity can help language learners develop the authentic use of the language and drill real-life interaction skills (J. Han & Lee, 2024). It is also necessary to advance teaching approaches and curricula to effectively integrate AI chatbots into language education. A particular method could be to provide learning modules on various sorts of chatbots (i.e. mobile or web-based, etc.) in the curriculum. This could afford language learners and teachers a more intricate understanding of these tools and their possible ways of utilization in language education (Belda-Medina & Kokošková, 2023).

In addition, for the chatbots to function as good conversation partners, there is a need for the AI system to converse on a broader range of topics. The latest developments in AI that could contribute to this aspect are language models such as Google's LaMDA and OpenAI's GPT-4. Founded on the extension of large language models, they promise to merge the benefits of maintaining an open-ended interaction with characteristics that could elevate learners' language learning benefits and motivation (Godwin-Jones, 2022). The scope of further studies on voice-based chatbots can focus on such language models in different language learning settings and with different student profiles.

5. Conclusion

The results of this meta-synthesis have demonstrated that chatbots have gained noticeable consideration in English language learning research. The results of this review corroborated the positive impacts of employing voice-based chatbots on English language learning. Based on our findings, implications for teachers and researchers could be proposed. Teachers and researchers should reevaluate the use of voice-based chatbots in English language learning research, as they are a potentially valuable means of developing listening and speaking skills, vocabulary knowledge, and pragmatic and oral communicative competence.

The meta-synthesis findings suggest that the international literature on voice-based AI chatbot use in the English language field needs more varied methods, theoretical frameworks, and research trends. Our findings may contribute to the line of knowledge on this subject and address these challenges as well as potential affordances for upcoming studies. Practitioners and researchers can employ the current research findings to further the practical uses of dynamically evolving voiced-based chatbots in language classes.

Supplementary material. To view supplementary material referred to in this article, please visit https://doi.org/10.1017/S0958344024000168

Ethical statement and competing interests. This study did not involve any human or animal participants, and thus no ethical approval was required. All the studies included in this meta-synthesis research can be accessed online. The authors declare no competing interests. The authors also declare no use of generative AI.

References

*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: Article 100148. https://doi.org/10.1016/j.caeai.2023.100148

https://doi.org/10.1017/S0958344024000168 Published online by Cambridge University Press

*Ayedoun, E., Hayashi, Y. & Seta, K. (2019a) Adding communicative and affective strategies to an embodied conversational agent to enhance second language learners' willingness to communicate. *International Journal of Artificial Intelligence in Education*, 29(1): 29–57. https://doi.org/10.1007/s40593-018-0171-6

*Ayedoun, E., Hayashi, Y. & Seta, K. (2019b) L2 learners' preferences of dialogue agents: A key to achieve adaptive motivational support? In Isotani, S., Millán, E., Ogan, A., Hastings, P., McLaren, B. & Luckin, R. (eds.), *Artificial intelligence in education: 20th International Conference, AIED 2019, Chicago, IL, USA, June 25-20, 2019: Proceedings, Part II*. Cham: Springer International Publishing, 19–23.

*Ayedoun, E., Hayashi, Y. & Seta, K. (2020) Toward personalized scaffolding and fading of motivational support in L2 learner–dialogue agent interactions: An exploratory study. *IEEE Transactions on Learning Technologies*, 13(3): 604–616. https://doi.org/10.1109/TLT.2020.2989776

*Belda-Medina, J. & Calvo-Ferrer, J. R. (2022) Using chatbots as AI conversational partners in language learning. *Applied Sciences*, 12(17): Article 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: Article 62. https://doi.org/10.1186/s41239-023-00432-3

*Çakmak, F. (2022) Chatbot-human interaction and its effects on EFL students' L2 speaking performance and anxiety. *Novitas-ROYAL*, 16(2): 113–131.

Caldarini, G., Jaf, S. & McGarry, K. (2022) A literature survey of recent advances in chatbots. *Information*, 13(1): Article 41. https://doi.org/10.3390/INFO13010041

*Chen, H. H.-J., Yang, C. T.-Y. & Lai, K. K.-W. (2023) Investigating college EFL learners' perceptions toward the use of Google Assistant for foreign language learning. *Interactive Learning Environments*, 31(3): 1335–1350. https://doi.org/10.1080/10494820.2020.1833043

Davis, F. D. (1989) Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 13(3): 319–340. https://doi.org/10.2307/249008

*Dizon, G. (2017) Using intelligent personal assistants for second language learning: A case study of Alexa. *TESOL Journal*, 8(4): 811–830. https://doi.org/10.1002/tesj.353

*Dizon, G. (2020) Evaluating intelligent personal assistants for L2 listening and speaking development. *Language Learning & Technology*, 24(1): 16–26. https://doi.org/10125/44705

*Dizon, G. & Tang, D. (2020) Intelligent personal assistants for autonomous second language learning: An investigation of Alexa. *The JALT CALL Journal*, 16(2): 107–120. https://doi.org/10.29140/jaltcall.v16n2.273

*Ebadi, S. & Amini, A. (2024) Examining the roles of social presence and human-likeness on Iranian EFL learners' motivation using artificial intelligence technology: A case of CSIEC chatbot. *Interactive Learning Environments*, 32(2): 655–673. https://doi.org/10.1080/10494820.2022.2096638

Ellis, R. (2003) *Task-based language learning and teaching*. Oxford: Oxford University Press.

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

Epley, N., Waytz, A. & Cacioppo, J. T. (2007) On seeing human: A three-factor theory of anthropomorphism. *Psychological Review*, 114(4): 864–886. https://doi.org/10.1037/0033-295X.114.4.864

*Ericsson, E., Lundin, J. & Sofkova Hashemi, S. (2023) From deadpan machine to relating socially: Middle school students' experiences speaking English with embodied conversational agents. *Journal of Research on Technology in Education*. Advance online publication. https://doi.org/10.1080/15391523.2023.2213902

*Ericsson, E., Sofkova Hashemi, S. & Lundin, J. (2023) Fun and frustrating: Students' perspectives on practicing speaking English with virtual humans. *Cogent Education*, 10(1): Article 2170088. https://doi.org/10.1080/2331186X.2023.2170088

Finfgeld, D. L. (2003) Metasynthesis: The state of the art—So far. *Qualitative Health Research*, 13(7): 893–904. https://doi.org/10.1177/1049732303253462

Fishbein, M. & Ajzen, I. (1975) *Belief, attitude, intention, and behavior: An introduction to theory and research*. Reading: Addison-Wesley.

*Forsyth, C. M., Luce, C., Zapata-Rivera, D., Jackson, G. T., Evanini, K. & So, Y. (2019) Evaluating English language learners' conversations: Man vs. machine. *Computer Assisted Language Learning*, 32(4): 398–417. https://doi.org/10.1080/09588221.2018.1517126

*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. https://doi.org/10.1016/j.chb.2017.05.045

*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

*Fryer, L. K., Thompson, A., Nakao, K., Howarth, M. & Gallacher, A. (2020) Supporting self-efficacy beliefs and interest as educational inputs and outcomes: Framing AI and human partnered task experiences. *Learning and Individual Differences*, 80: Article 101850. https://doi.org/10.1016/j.lindif.2020.101850

Gardner, D. & Miller, L. (1999) *Establishing self-access: From theory to practice*. Cambridge: Cambridge University Press.

https://doi.org/10.1017/S0958344024000168 Published online by Cambridge University Press

Garrison, D. R. & Arbaugh, J. B. (2007) Researching the community of inquiry framework: Review, issues, and future directions. *The Internet and Higher Education*, 10(3): 157–172. https://doi.org/10.1016/j.iheduc.2007.04.001

Gibson, J. J. (1986) *The ecological approach to visual perception*. Hillsdale: Lawrence Erlbaum.

Godwin-Jones, R. (2022) Chatbots in language learning: AI systems on the rise. In Arnbjörnsdóttir, B., Bédi, B., Bradley, L., Friðriksdóttir, K., Garðarsdóttir, H., Thouësny, S. & Whelpton, M. J. (eds.), *Intelligent CALL, granular systems and learner data: Short papers from EUROCALL 2022*. Voillans: Research-publishing.net, 124–128. https://doi.org/10.14705/rpnet.2022.61.1446

Godwin-Jones, R. (2023) Smart devices and informal language learning. In Toffoli, D., Sockett, G. & Kusyk, M. (eds.), *Language learning and leisure: Informal language learning in the digital age*. Berlin: Walter de Gruyter GmbH, 89–113. https://doi.org/10.1515/9783110752441-004

*Han, D.-E. (2020) The effects of voice-based AI chatbots on Korean EFL middle school students' speaking competence and affective domains. *Asia-Pacific Journal of Convergent Research Interchange*, 6(7): 71–80. https://doi.org/10.47116/apjcri.2020.07.07

Han, J. & Lee, D. (2024) Research on the development of principles for designing elementary English speaking lessons using artificial intelligence chatbots. *Humanities and Social Sciences Communications*, 11: Article 212. https://doi.org/10.1057/s41599-024-02646-w

*Hassani, K., Nahvi, A. & Ahmadi, A. (2016) Design and implementation of an intelligent virtual environment for improving speaking and listening skills. *Interactive Learning Environments*, 24(1): 252–271. https://doi.org/10.1080/10494820.2013.846265

Hidi, S. & Renninger, K. A. (2006) The four-phase model of interest development. *Educational Psychologist*, 41(2): 111–127. https://doi.org/10.1207/s15326985ep4102_4

Horwitz, E. K., Horwitz, M. B. & Cope, J. (1986) Foreign language classroom anxiety. *The Modern Language Journal*, 70(2): 125–132. https://doi.org/10.2307/327317

*Hsu, L. (2022) To CALL or not to CALL: Empirical evidence from neuroscience. *Computer Assisted Language Learning*, 35(4): 792–815. https://doi.org/10.1080/09588221.2020.1750429

*Hsu, M.-H., Chen, P.-S. & Yu, C.-S. (2023) Proposing a task-oriented chatbot system for EFL learners speaking practice. *Interactive Learning Environments*, 31(7): 4297–4308. https://doi.org/10.1080/10494820.2021.1960864

Huang, W., Hew, K. F. & Fryer, L. K. (2022) Chatbots for language learning—Are they really useful? A systematic review of chatbot-supported language learning. *Journal of Computer Assisted Learning*, 38(1): 237–257. https://doi.org/10.1111/jcal.12610

*Hwang, W.-Y., Guo, B.-C., Hoang, A., Chang, C.-C. & Wu, N.-T. (2022) Facilitating authentic contextual EFL speaking and conversation with smart mechanisms and investigating its influence on learning achievements. *Computer Assisted Language Learning*. Advance online publication. https://doi.org/10.1080/09588221.2022.2095406

*Jeon, J. (2022) Exploring a self-directed interactive app for informal EFL learning: A self-determination theory perspective. *Education and Information Technologies*, 27(4): 5767–5787. https://doi.org/10.1007/s10639-021-10839-y

*Jeon, J. (2024) Exploring AI chatbot affordances in the EFL classroom: Young learners' experiences and perspectives. *Computer Assisted Language Learning*, 37(1–2): 1–26. https://doi.org/10.1080/09588221.2021.2021241

Jeon, J., Lee, S. & Choe, H. (2023) Beyond ChatGPT: A conceptual framework and systematic review of speech-recognition chatbots for language learning. *Computers & Education*, 206: Article 104898. https://doi.org/10.1016/j.compedu.2023.104898

Jeon, J., Lee, S. & Choi, S. (2023) A systematic review of research on speech-recognition chatbots for language learning: Implications for future directions in the era of large language models. *Interactive Learning Environments*. Advance online publication. https://doi.org/10.1080/10494820.2023.2204343

Ji, H., Han, I. & Ko, Y. (2023) A systematic review of conversational AI in language education: Focusing on the collaboration with human teachers. *Journal of Research on Technology in Education*, 55(1): 48–63. https://doi.org/10.1080/15391523.2022.2142873

Jia, J. (2009) CSIEC: A computer assisted English learning chatbot based on textual knowledge and reasoning. *Knowledge-Based Systems*, 22(4): 249–255. https://doi.org/10.1016/j.knosys.2008.09.001

*Johnson, W. L. (2019) Data-driven development and evaluation of Enskill English. *International Journal of Artificial Intelligence in Education*, 29(3): 425–457. https://doi.org/10.1007/s40593-019-00182-2

Jonassen, D., Davidson, M., Collins, M., Campbell, J. & Haag, B. B. (1995) Constructivism and computer-mediated communication in distance education. *American Journal of Distance Education*, 9(2): 7–26. https://doi.org/10.1080/08923649509526885

*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, N.-Y. (2018) A study on chatbots for developing Korean college students' English listening and reading skills. *Journal of Digital Convergence*, 16(8): 19–26. https://doi.org/10.14400/JDC.2018.16.8.019

Kim, N.-Y., Cha, Y. & Kim, H.-S. (2019) Future English learning: Chatbots and artificial intelligence. *Multimedia-Assisted Language Learning*, 22(3): 32–53. https://doi.org/10.15702/mall.2019.22.3.32

https://doi.org/10.1017/S0958344024000168 Published online by Cambridge University Press

Kukulska-Hulme, A. & Lee, H. (2020) Intelligent assistants in language learning: An analysis of features and limitations. In Frederiksen, K.-M., Larsen, S., Bradley, L. & Thouësny, S. (eds.), *CALL for widening participation: Short papers from EUROCALL 2020*. Voillans: Research-publishing.net, 172–176. https://doi.org/10.14705/rpnet.2020.48.1184

Lantolf, J. P., Thorne, S. L. & Poehner, M. E. (2014) Sociocultural theory and second language development. In VanPatten, B. & Williams, J. (eds.), *Theories in second language acquisition: An introduction* (2nd ed.). New York: Routledge, 207–226.

*Lee, K.-A. & Lim, S.-B. (2023) Designing a leveled conversational teachable agent for English language learners. *Applied Sciences*, 13(11): Article 6541. https://doi.org/10.3390/app13116541

*Lee, S. & Jeon, J. (2022) Visualizing a disembodied agent: Young EFL learners' perceptions of voice-controlled conversational agents as language partners. *Computer Assisted Language Learning*. Advance online publication. https://doi.org/10.1080/09588221.2022.2067182

*Li, K.-C., Chang, M. & Wu, K.-H. (2020) Developing a task-based dialogue system for English language learning. *Education Sciences*, 10(11): Article 306. https://doi.org/10.3390/educsci10110306

Long, M. H. (1996) The role of the linguistic environment in second language acquisition. In Ritchie, W. C. & Bhatia, T. K. (eds.), *Handbook of second language acquisition*. New York: Academic Press, 413–468. https://doi.org/10.1016/B978-012589042-7/50015-3

Lyster, R. & Sato, M. (2013) Skill acquisition theory and the role of practice in L2 development. In García Mayo, M. P., Gutierrez Mangado, M. J. & Martínez Adrián, M. (eds.), *Contemporary approaches to second language acquisition*. Amsterdam: John Benjamins, 71–92. https://doi.org/10.1075/aals.9.07ch4

MacIntyre, P. D., Clément, R., Dörnyei, Z. & Noels, K. A. (1998) Conceptualizing willingness to communicate in a L2: A situational model of L2 confidence and affiliation. *The Modern Language Journal*, 82(4): 545–562. https://doi.org/10.1111/j.1540-4781.1998.tb05543.x

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G. & the PRISMA Group. (2009) Reprint—Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *Physical Therapy*, 89(9): 873–880. https://doi.org/10.1093/ptj/89.9.873

*Morton, H., Gunson, N. & Jack, M. (2012) Interactive language learning through speech-enabled virtual scenarios. *Advances in Human-Computer Interaction*, 2012: Article 389523. https://doi.org/10.1155/2012/389523

*Morton, H. & Jack, M. (2010) Speech interactive computer-assisted language learning: A cross-cultural evaluation. *Computer Assisted Language Learning*, 23(4): 295–319. https://doi.org/10.1080/09588221.2010.493524

*Moussalli, S. & Cardoso, W. (2020) Intelligent personal assistants: Can they understand and be understood by accented L2 learners? *Computer Assisted Language Learning*, 33(8): 865–890. https://doi.org/10.1080/09588221.2019.1595664

*Moussalli, S. & Cardoso, W. (2021) Intelligent personal assistants and L2 pronunciation development: Focus on English past -ed. In Zoghlami, N., Brudermann, C., Sarré, C., Grosbois, M., Bradley, L. & Thouësny, S. (eds.), *CALL and professionalisation: Short papers from EUROCALL 2021*. Voillans: Research-publishing.net, 226–231. https://doi.org/10.14705/rpnet.2021.54.1337

Nass, C., Moon, Y., Fogg, B. J., Reeves, B. & Dryer, D. C. (1995) Can computer personalities be human personalities? *International Journal of Human-Computer Studies*, 43(2): 223–239. https://doi.org/10.1006/ijhc.1995.1042

*Park, H. (2022) Effects of virtual reality-based English learning on Korean university students' speaking ability. *Multimedia-Assisted Language Learning*, 25(4): 93–119. https://doi.org/10.15702/mall.2022.25.4.93

Pérez, J. Q., Daradoumis, T. & Puig, J. M. M. (2020) Rediscovering the use of chatbots in education: A systematic literature review. *Computer Applications in Engineering Education*, 28(6): 1549–1565. https://doi.org/10.1002/cae.22326

Rassaei, E. (2023) The effects of text-based and audio-based dynamic glosses on L2 vocabulary learning: A dynamic assessment approach. *The Language Learning Journal*, 51(4): 509–522. https://doi.org/10.1080/09571736.2023.2213247

Reinders, H. & White, C. (2016) 20 years of autonomy and technology: How far have we come and where to next? *Language Learning & Technology*, 20(2): 143–154. https://doi.org/10125/44466

Ryan, R. M. & Deci, E. L. (2017) *Self-determination theory: Basic psychological needs in motivation, development, and wellness*. New York: Guilford Press. https://doi.org/10.1521/978.14625/28806

Saini, M. & Shlonsky, A. (2012) *Systematic synthesis of qualitative research*. New York: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195387216.001.0001

Schmidt, R. (1995) Consciousness and foreign language learning: A tutorial on the role of attention and awareness in learning. In Schmidt, R. (ed.), *Attention and awareness in foreign language learning*. Honolulu: University of Hawai'i Mānoa, Second Language Teaching & Curriculum Center, 1–63.

*Sydorenko, T., Daurio, P. & Thorne, S. L. (2018) Refining pragmatically-appropriate oral communication via computer-simulated conversations. *Computer Assisted Language Learning*, 31(1–2): 157–180. https://doi.org/10.1080/09588221.2017.1394326

*Sydorenko, T., Smits, T. F. H., Evanini, K. & Ramanarayanan, V. (2019) Simulated speaking environments for language learning: Insights from three cases. *Computer Assisted Language Learning*, 32(1–2): 17–48. https://doi.org/10.1080/09588221.2018.1466811

https://doi.org/10.1017/S0958344024000168 Published online by Cambridge University Press

*Tai, T.-Y. & Chen, H. H.-J. (2022) The impact of intelligent personal assistants on adolescent EFL learners' listening comprehension. *Computer Assisted Language Learning*. Advance online article. https://doi.org/10.1080/09588221.2022.2040536

*Tai, T.-Y. & Chen, H. H.-J. (2023) The impact of Google Assistant on adolescent EFL learners' willingness to communicate. *Interactive Learning Environments*, 31(3): 1485–1502. https://doi.org/10.1080/10494820.2020.1841801

Thornberg, R. & Charmaz, K. (2014) Grounded theory and theoretical coding. In Flick, U. (ed.), *The SAGE handbook of qualitative data analysis*. London: SAGE Publications, 153–169. https://doi.org/10.4135/9781446282243.n11

*Timpe-Laughlin, V. & Dombi, J. (2020) Exploring L2 learners' request behavior in a multi-turn conversation with a fully automated agent. *Intercultural Pragmatics*, 17(2): 221–257. https://doi.org/10.1515/ip-2020-0010

*Timpe-Laughlin, V., Sydorenko, T. & Daurio, P. (2022) Using spoken dialogue technology for L2 speaking practice: What do teachers think? *Computer Assisted Language Learning*, 35(5–6), 1194–1217. https://doi.org/10.1080/09588221.2020.1774904

*Wang, X., Liu, Q., Pang, H., Tan, S. C., Lei, J., Wallace, M. P. & Li, L. (2023) What matters in AI-supported learning: A study of human-AI interactions in language learning using cluster analysis and epistemic network analysis. *Computers & Education*, 194: Article 104703. https://doi.org/10.1016/j.compedu.2022.104703

*Wang, X., Pang, H., Wallace, M. P., Wang, Q. & Chen, W. (2022) Learners' perceived AI presences in AI-supported language learning: A study of AI as a humanized agent from community of inquiry. *Computer Assisted Language Learning*. Advance online article. https://doi.org/10.1080/09588221.2022.2056203

*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: 1–22. https://doi.org/10.1177/18479790231176372

Xiao, F., Zhao, P., Sha, H., Yang, D. & Warschauer, M. (2023) Conversational agents in language learning. *Journal of China Computer-Assisted Language Learning*. Advance online publication. https://doi.org/10.1515/jccall-2022-0032

*Yang, C. T.-Y., Lai, S.-L. & Chen, H. H.-J. (2022) The impact of intelligent personal assistants on learners' autonomous learning of second language listening and speaking. *Interactive Learning Environments*. Advance online publication. https://doi.org/10.1080/10494820.2022.2141266

*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. https://doi.org/10.1017/S0958344022000039

Yuan, Y. (2023) An empirical study of the efficacy of AI chatbots for English as a foreign language learning in primary education. *Interactive Learning Environments*. Advance online publication. https://doi.org/10.1080/10494820.2023.2282112

Zhai, C. & Wibowo, S. (2022) A systematic review on cross-culture, humor and empathy dimensions in conversational chatbots: The case of second language acquisition. *Heliyon*, 8(12): Article e12056. https://doi.org/10.1016/j.heliyon.2022.e12056

About the authors

Fatma Şeyma Koç is currently a research assistant in the Department of English Language Teaching at Akdeniz University, Türkiye. She holds a PhD in the English Language Teaching program at Middle East Technical University. Her research interests include teaching language skills via instructional technology in English as a foreign language classes.

Perihan Savaş is a professor of English as a foreign language (EFL) at Middle East Technical University, Türkiye. She received her PhD from the University of Florida. Her scholarly interests include integrating technology into EFL curriculum, mobile-assisted language learning (MALL), AI in EFL, and teacher training/faculty support in online education.

https://doi.org/10.1017/S0958344024000168 Published online by Cambridge University Press