A Systematic Review of Voice-based Intelligent Virtual Agents in EFL Education

Eirene Katsarou[1(✉)], Fridolin Wild[2], Areti-Maria Sougari[3], Paraskevi Chatzipanagiotou[4]
[1]Democritus University of Thrace, Komotini, Greece
[2]The Open University, Milton Keynes, UK
[3]Aristotle University of Thessaloniki, Thessaloniki, Greece
[4]European University, Engomi, Cyprus
ekatsaro@fmenr.duth.gr

Abstract—Since its debut in the field of education nearly three decades ago, Artificial Intelligence (AI) has been considered a powerful tool to facilitate new paradigms for instructional design and innovative educational practice by means of intelligent tutoring systems, adaptive learning systems, educational chatbots, teaching robots, to name a few. Recent technological advances in the adjacent areas of natural language processing, machine learning, and computer graphics focusing primarily on design features that can improve their human-like qualities of naturalness and believability as human interlocutors have also amplified new application opportunities for Intelligent Virtual Agents (IVAs) or Animated Pedagogical Agents (APAs) within the area of Intelligent Computer-Assisted Language Learning (ICALL). Although AI-powered IVAs hold the potential to improve the learning process in nearly any knowledge domain and personalize automation in teaching by embodying different roles in the learning environment, strikingly few studies have empirically attempted to assess IVAs impact on L2 learners' academic achievement when learning English as a Foreign (EFL) so far. The present article addresses this issue via a systematic review of relevant interventionist IVA studies that were conducted in EFL settings and published within the 2015–2020 timeframe examining IVAs key affordances, major barriers in their adoption for language learning purposes, and the CALL research trends currently prevalent on the topic. Pedagogical implications for the effective implementation of IVA technology in L2 contexts are discussed and future research avenues in the area are highlighted.

Keywords—intelligent virtual agents, EFL, L2 education

1 Introduction

The advent of Artificial Intelligence (AI), as an emerging technology in education, has been acknowledged to hold the potential to revolutionize educational practice and foster a transformation of traditional educational systems via task automation and mechanization [32]. This has been manifested in the gradual proliferation of various

applications intended to provide customized learning [33], [46], [21], offer dynamic assessments, and facilitate meaningful interactions in online or blended learning experiences [74]. This is also reflected in Baker et al.'s [4] recent classification of educational AI (AIEd) tools into instructional-oriented systems, used by learners to learn a subject matter, i.e. *learner-facing* and *teacher-facing*, used by teachers to reduce their workload and obtain insight into learners' learning progress to proactively and promptly offer feedback and guidance in the process, while *system-facing* AI systems provide information for administrators and managers on the institutional level. Intelligent tutoring systems (ITS), virtual agents, and intelligent virtual reality have been cited as the most commonly available AI-based learning systems currently in use to support personalized and collaborative modes of learning [45]. This trend is also prevalent in ICALL contexts where a substantial amount of extant empirical research has predominantly been dedicated to an investigation of the pedagogical expediency of chatbot- [20], [35] and ITS-assisted approaches [13] in foreign language learning to the exception of Intelligent Virtual Agents (IVAs) up until recently.

In practice, the development of IVAs materializes, in theory, the notion of human-like thinking in technological terms [61], performing various tasks with several challenges in diverse sectors of human life, yet IVAs have often been defined rather inconsistently by experts in computer engineering. Graesser & McNamara [27] describe IVAs as either scripted or intelligently adaptive learning environments with talking heads and facial expressions that speak, point, gesture, and instruct the learner what to do [37] as experts, mentors [40], tutors [54] and learning companions [24]. In Rickel et al. [59], IVAs act as autonomous agents, immersed in a virtual world, capable of face-to-face interaction, equipped with a human face, and synchronized lip motions with dialog. Humanoid embodiment and IVAs cognitive capabilities are highlighted by Traum [69] as key traits of IVAs that enable users' engagement in human activities and meaningful interaction. This contradicts Burden and Savin-Baden's [9] most recent definition where a humanoid body is not considered an essential characteristic of IVAs "as long as they exhibit human-like behaviors (speech, gesture, and movement) and other human characteristics (emotions, empathy, reasoning, planning, motivation) and can be seen on a computer screen, heard through a speaker, or accessed in some other way" [9, p.13]. The all-encompassing nature of this definition suits the purposes of this review as it seeks to address the pedagogical potential of voice-based, general-use IVAs (Amazon Alexa, Siri, Google Assistant) in L2 learning practice. Based on empirical research evidence available in relevant interventionist IVA studies in EFL settings, this review critically assesses the role of AI in the EFL classroom by exploring IVA integration as an emerging trend within L2 teaching practice and learning based on the following research questions:

RQ#1. What are the pedagogical affordances of IVAs in L2 educational settings?
RQ#2. What are the barriers inhibiting the adoption of IVAs for EFL learning?
RQ#3. What are the current research trends followed with respect to IVA use in the EFL context?

The review is organized as follows: Section 2 focuses on the nature of IVA technology as distinguished from related AI-supported learning systems in the field of education

followed by an outline of the study's research methodology in Section 3. Review results are discussed per research question in Section 4 and are further summarized in Section 5 along with implications and future research avenues for EFL practitioners and CALL instructional designers while concluding remarks appear in the last section.

2   Types of AI-based virtual agents

When attempting to define IVAs, it is useful to examine the different traits that could be used to differentiate them from other forms of computerized systems that may either show intelligence or are based on some human characteristics. As illustrated by previous research on AIEd technology [5], [34], this becomes a necessity as the boundaries between what is an IVA and what is not are rather fuzzy, and AIEd is often applied invariably to denote a wide spectrum of AI entities such as chatbots, teachable agents (ITS), embodied conversational agents (ECAs) and tutors. Although similar to IVAs' design and philosophy, these AI-supported systems differ technologically, in terms of type of intelligence and humanoid features they possess, to effectively sustain dialogue-based CALL leading to more positive social interactions, and thus allowing learners to practice in L2 meaningful conversations autonomously [6] as illustrated below.

Chatbots, also known as bots or interactive agents, are computer applications that respond like smart entities when conversed with through text or voice [39] due to embedded Natural Language Processing (NLP) technology [2], offering wide applications in education, health care, and marketing [1] as platform-independent tools that are instantly available to users without needed installations. When a chatbot is able to communicate with a user, it acts as an advisor and a companion with whom the user can build a long-term relationship [57], as is the case of Wysa, an AI-based commercial smartphone application in which the user can chat with to reduce stress by texting without exploiting non-verbal behaviors. In the field of CALL, chatbot technology has been found to be associated in practice with an increase in L2 learning motivation [44], an improvement of overall L2 proficiency [14] as well as with minor facilitation of specific language skills such as lexical inferencing [36], while offering ample conversation opportunities for practice [68]. On the other hand, *intelligent tutoring systems* (ITSs) or teachable agents [48] are computer learning systems intended to detect complex principles of learning and to help learners acquire declarative knowledge and procedural skills based on powerful intelligent algorithms adjusted to learners' educational needs [25]. ITS systems work with one student at a time. They track learners' knowledge, skills, idiosyncratic profiles of cognitive and affective attributes, as well as other psychological traits [62] and adaptively respond by utilizing computational mechanisms in artificial intelligence and cognitive science [26]. Although cognitively intelligent, both chatbots and ITSs are clearly differentiated from IVAs in view of the absence of human-like traits and as such they are examples of an Artificial Narrow Intelligence (ANI) stage (Figure 1).

Embodied conversational agents (ECAs), defined in 2000 by Justine Cassell [10], replaced the textual interaction of chatbots and the voice-only interaction of intelligent voice assistants with a more natural interaction, combining verbal and non-verbal communication; their presence seems to improve the user's interaction with this

AI system. ECAs are, in fact, endowed with a humanoid body, capable of exploiting human-specific communication modalities, such as voice and facial expression, gaze, gestures, head movements, posture, and displacement [7]. They represent a more intuitive interface between the user and the computer system, and their level of complexity can vary greatly depending on the context and applications in which they are deployed. Still mainly present in the world of ASI research, they are often used to validate psychological theories of human behavior by concretizing them through computer models that control and generate the behavior of the virtual agent [66]. *Tutors* also form a sub-class of embodied agents exhibiting an increased level of complexity used mainly for educational purposes provided that they have a knowledge base in a certain field as well as the ability to help the learner by providing easy access to information and boosting motivation.

Intelligent Virtual Assistants (IVAs), or intelligent personal assistants (IPAs) can be distinguished from other AI expert systems in that they are restricted to user assistance. They represent autonomous entities capable of perceiving their environment and acting on it to achieve a goal [60] by interacting with humans mainly through a synthetic voice and assisting users to perform generic tasks to improve their daily lives [19]. To understand the user, they use automatic natural language processing to match the user's text with executable commands [67] while many IVAs learn continuously using artificial intelligence (AI) techniques, including machine learning. A virtual assistant (incarnated or not) differs from a virtual companion as, contrary to the latter, it is neither programmed to hold long conversations nor to create relationships. However, the virtual assistant can exhibit characteristics of the virtual companion, such as a personality to better perform their tasks, offer a more convincing experience, and imitate human relationships [56]. The explicit goal of a virtual assistant is to ensure natural and inherently cognitive, linguistic, and collaborative interactions with the user undertaken in a fluid manner akin to meaningful and mutually reciprocal human communication. Amazon's Alexa, Apple's Siri, and Google Assistant are the three predominant cloud-based, general-use IVAs in the world of artificial intelligence, widely used in mobile and stationary devices alike for swift and convenient two-way communication with users through hands-free control and verbal responses [17]. These assist in completing basic tasks such as consulting the weather forecast, checking the latest news, and setting reminders [71]. Moreover, persona voice assistants' search results in terms of language and location are more relevant to the user [18]. Following the work by Burden & Savin-Baden [9], the main characteristics of IVAs are as presented below [9, p.14]:

- Manifests itself in a visual, auditory, textual, or similar form,
- May have some embodiment within a virtual world,
- May present itself as humanoid in manifestation and behavior,
- Will have a natural language capability,
- Must exhibit a degree of autonomy,
- May have the ability to express, recognize and respond to emotions
- May exhibit some aspects of a personality,
- May have some ability to reason in a human-like way,
- May exhibit some elements of imagination, and
- May even have a self-narrative, but unlikely to have any indications of sentience.

Giles and Bevacqua [22] add the physical medium to this list, i.e., the agent can "inhabit" a single device or migrate into several devices.

Based on the type of intelligence (cognitive, emotional, and social) displayed, IVAs can further be classified into analytical, human-inspired, and humanized, representing different stages in the evolution of AI-supported IVAs technology as suggested by Kaplan & Haelein [38].
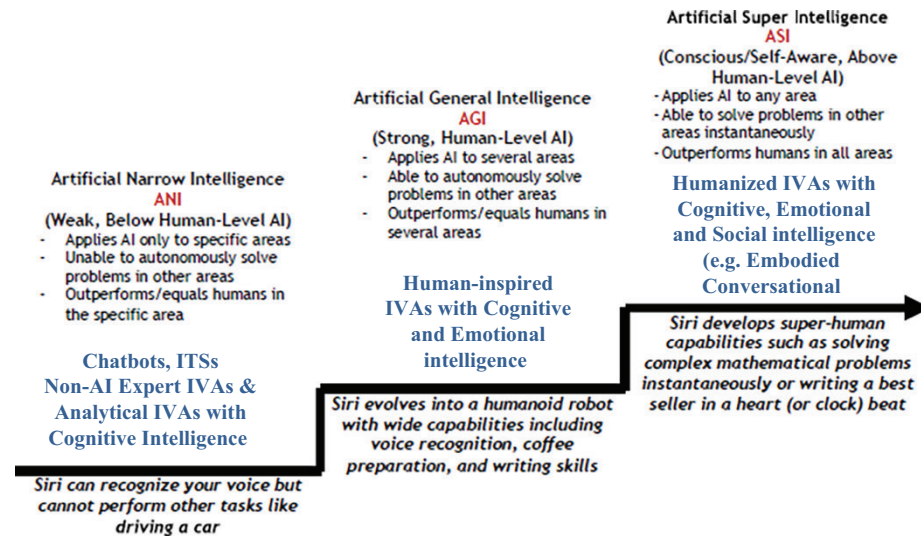

Fig. 1. Evolution of AI technology (Adapted from: Kaplan & Haenlein, 2019)

Expert Virtual Assistants belong to first-generation AI systems and have no kind of intelligence, except for the inherent information program it is endowed with. The program uses "If – Then" algorithm to complete tasks and can only answer certain questions [38, p. 18]. When interacting with users, they only offer installed alternatives to questions without any additional clarifying information while emotion recognition and customer interaction through word search are also not feasible options. Compared to Expert Virtual Assistants, Analytical Virtual Assistants use cognitive intelligence. They analyze past experiences to make future decisions and generate a cognitive representation similar to the real world. In turn, *Human-Inspired Virtual Assistants* differ from the Analytical one in the sense that they also possess emotional intelligence; this greatly affects their decision-making processes. They are mostly human-like characters, acting as attentive listeners and effective interlocutors and displaying life-like behaviors such as speech, locomotion, gestures, and facial expressions (i.e. in line with Burden& Savin-Baden's [9] characteristics outlined above. They express reactions and make a conversation based on the emotions of a person, as well as counterfeit human emotions [12]. *Humanized Virtual Assistants* are still evolving and are considered the most sophisticated, combining cognitive, emotional, and social intelligence in one. In the near future, they are envisaged as the main tool to hold and analyze considerable amounts of past experience that will enhance interaction with people and probably with

other assistants without human effort [47], while at the same time exhibiting additional features of self-consciousness [50].

In contrast to other AI software applications, the innovative element of IVAs in education hangs on two fundamental aspects that critically affect how they offer vigorous scaffoldings and easily provide additional help to learners with advanced technical voice tools [51], i.e., (i) using underlying technology based on deep learning technology to recognize learners' utterances without necessitating self-generated training data [43] and (ii) integrating IVAs in devices (e.g., Google's assistant, Apple's Siri on smartphones, and Microsoft's Cortana on desktop PCs), ascertaining ease of accessibility mostly through clicking or giving spoken commands that can render IVAs as daily aides in everyday affairs. To what extent the potential of IVA implementation, as an emerging technology can be extended and maximized in the educational practice of L2 learning remains to be examined within the realm of this review study.

3 Methodology for the review

This study adopts a systematic approach to reviewing the literature on intelligent virtual agents to obtain comprehensive insights into the state-of-the-art [41].

The scientific method of systematic research is considered to be superior when compared to conventional literature reviews as it enhances certain aspects such as consistency, replicability, reliability, and validity [73] and reduces redundancies in the published literature, allowing disparities and trends for future studies to be identified. The review process followed here is conducted within the 'Preferred Reporting Items for Systematic Reviews and Meta-Analyses' (PRISMA) framework and can be divided into three concrete steps [42]:

Phase 1 *Planning* includes journal selection, delineation of inclusion and exclusion criteria for study selection, and definition of categories utilized in the analysis.

Phase 2 *Conducting the review* involves study selection, data extraction, synthesis, and coding scheme.

Phase 3 *Reporting the review* consists of result analysis and discussion of main results, tendencies, implications, and conclusions.

To further elucidate the procedure followed, the following stages were deemed necessary to consider:

3.1 Search strategy

This study reviewed only relevant work published in English on IVAs within the timeframe spanning from 2015 up until 2022. *Scopus*, *ScienceDirect*, *Google Scholar,* and *CrossRef* were the primary databases used to serve the purposes of our study and were searched separately. Highly-esteemed international peer-reviewed journals in the fields of educational technology and CALL including *Computers & Education, Computers in Human Behavior Computers & Education: Artificial Intelligence*, *RECALL, CALICO, Computer-Assisted Language Learning, JALT CALL Journal, Language*

Learning & Technology were also manually searched. The keywords used in this review that helped the authors determine the scope and nature of virtual assistants in relation to education and training are: 'virtual humans', 'intelligent virtual assistants AND education/training', 'intelligent virtual agents AND education/training', and 'pedagogical agents'.

3.2 Selection criteria

To answer the research questions, a set of selection and quality criteria were determined to enable us to identify relevant empirical studies on the topic of IVAs in EFL education. The selection of the reviewed studies considered the following inclusion and exclusion criteria:

Inclusion Criteria	Exclusion Criteria
• Published between 2015 and the third quarter of 2022. • Written in English. • Availability of empirical evidence on IVAs educational potential, employing sound experimental research methodology conducted within an EFL context. • Studies that recounted the potential gains and challenges posed by IVAs adoption when implemented in EFL settings. • Studies that sufficiently illustrated an employed prototype of IVAs applied in EFL contexts, providing details on the artificial intelligence and machine learning techniques IVAs use to effectively help L2 learners learn.	• Not published journals and conference proceedings (e.g., books, reports, BA, MA, and PhD theses). • Not written in English. • Reviews, meta-analyses or overviews • Short reports that did not provide well-thought-out research methodology (e.g., studies in progress). • Studies that did not clearly summarize or aggregate the results of their research of qualitative and/or quantitative data.

3.3 Study quality assessment

Following the application of the inclusion and exclusion criteria, a checklist for the evaluation of the quality of the chosen articles from a methodological-design perspective was completed. Thus, emphasis was placed on empirically grounded analytical studies, which are considered to be the most accurate forms of experimental research to support or refute a hypothesis [58].

The educational potential of the reviewed studies was put under the lens based on the following aspects:

(a) the instructional system design and research methods employed in the studies to measure the effective implementation of IVAs within EFL contexts in terms of successful language learning outcomes.

(b) their scientific contribution with respect to the technological development and effective integration of AI-powered IVAs in EFL settings.

(c) the impact of IVAs use on EFL learners' academic performance, engagement and motivation when embedded to support different teaching approaches.

3.4 Data collection and data analysis

The publications that satisfied the aforementioned inclusion criteria were further categorized after considering past comparable systematic research [29], [31], [55] on the use of AI technology in education. Each one of the reviewed empirical studies served as the unit of analysis, while the coding system applied for data extraction arose inductively and was continuously improved through our interaction with the data [8]. Following the PRISMA principles, the literature search and selection process are presented in Figure 2 [49]. Ten (10) studies were found to be eligible for this review after duplicates were removed, abstracts were examined, and full-text papers were reviewed.
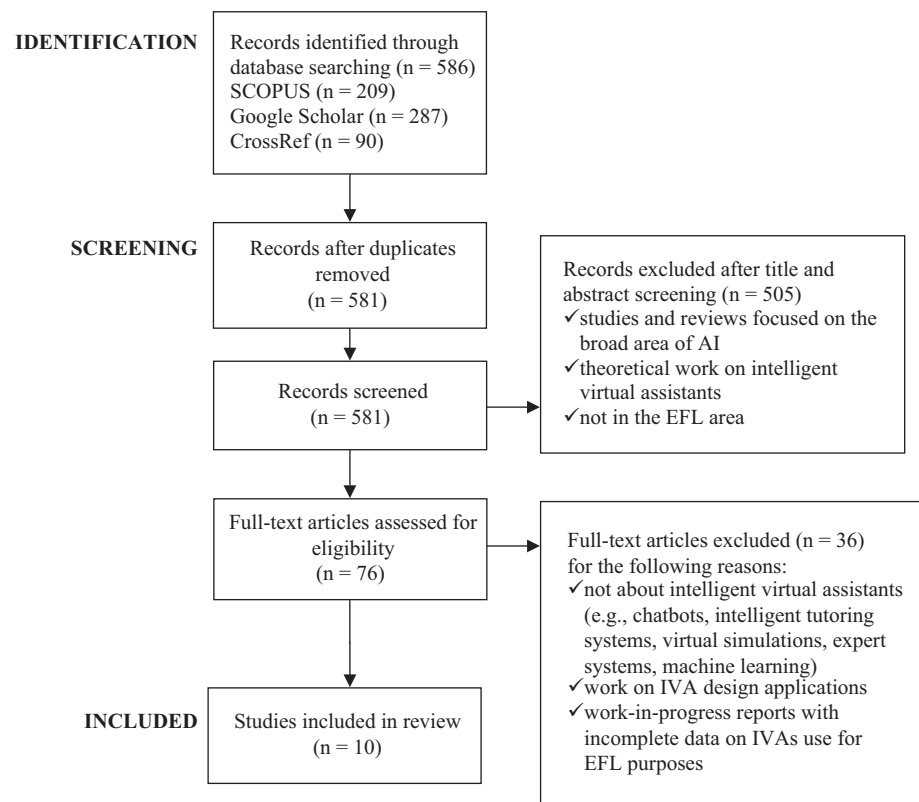

IDENTIFICATION — Records identified through database searching (n = 586)
SCOPUS (n = 209)
Google Scholar (n = 287)
CrossRef (n = 90)

SCREENING — Records after duplicates removed (n = 581)

Records screened (n = 581)

Records excluded after title and abstract screening (n = 505)
✓ studies and reviews focused on the broad area of AI
✓ theoretical work on intelligent virtual assistants
✓ not in the EFL area

Full-text articles assessed for eligibility (n = 76)

Full-text articles excluded (n = 36) for the following reasons:
✓ not about intelligent virtual assistants (e.g., chatbots, intelligent tutoring systems, virtual simulations, expert systems, machine learning)
✓ work on IVA design applications
✓ work-in-progress reports with incomplete data on IVAs use for EFL purposes

INCLUDED — Studies included in review (n = 10)

Fig. 2. The PRISMA process for literature search and selection

4 Results

4.1 Profile of selected studies

Prior work on the use of IVAs for L2 learning is scant and limited to a few ground-breaking, small-scale studies that have only recently spurred a renewed interest in the topic throughout the last five years. Table 1 displays a summary of all 10 studies considered in this review. All of the papers reported IVA implementation within a

formal instructional EFL context, with none being used in the digital wild. IVA application in the EFL classroom was explored in relation to its impact on different aspects of L2 language learning: L2 speaking and listening skills (7 studies), L2 learners' perceptions of IVA use (2 studies), and willingness to communicate (1 study). Results will be presented in this section to answer each of the initial research questions of this review.

4.2 Pedagogical affordances of IVAs in EFL

Underwood [70] is one of the earliest studies on the topic exploring the use of multiple voice-based IVAs (Alexa, Siri, and Google Assistant) with 11 elementary school EFL learners over a period of nine months, using a teacher-led design research study. Co-design strategies were further employed to encourage children to reflect on their IVA-enhanced L2 learning experiences and help them develop and express their own ideas about what AI language assistants might look like and how they might be used. Key findings of the study revealed that: (i) L2 learners' interactions with IVAs led to more meaningful L2 English exchanges and fun overall, even when miscommunication gaps occurred in cases when a virtual assistant failed to understand a particular command. Instead of giving up, the learners persisted and tended to rephrase their questions in ways more likely to be understood and answered by IVAs, (ii) although L2 learners faced some difficulty in understanding IVA responses due to their fast speech rate, L2 learners reported to have benefited most from their interactions with IVAs when aural input could also be displayed visually (e.g. use of Siri and Google Assistant) in smartphones and smart speakers with built-in displays.

In a follow-up study grounded on the interactionist approach to SLA and conducted within the tertiary education EFL context, Dizon [16] explored the potential of Amazon Alexa Echo Dot to support L2 listening and speaking skills for 37 undergraduate Japanese EFL beginner to intermediate learners in their first and second year of their studies who took the same elective English course to improve their communication skills through conversation, discussion, and presentation. Learners were initially surveyed in relation to their past experience with smart speakers responding that they had never used one prior to their participation in the quasi-experimental design of the study where the experimental group received a 10-week treatment of student-IVA interaction with Alexa, either individually or in pairs. Results showed that the treatment group significantly improved their L2 speaking proficiency, but not their L2 listening comprehension, as there was no significant difference between the control and experimental groups in this respect. This result was partly attributed either to EFL learners' inability to fully comprehend Alexa's responses as IVAs fail to successfully modify the output to promote enhanced L2 comprehension or to their focus on speaking practice during interactions, which may have thwarted them from paying close attention to Alexa's responses. Aligned with Underwood's [70] findings above, EFL learners' views of Alexa for in-class L2 learning were equally very positive, indicating that they not only enjoyed using the IVA for L2 learning but also perceived it to be a practical tool for learning English that could be utilized either for personalized or collaborative study. This finding correlates with Dizon and Tang's [15] mixed-method case design study on IVA use for self-directed, out-of-class language learning where EFL Japanese learners also perceived Alexa and the Echo Dot smart as a pedagogically effective tool with potential for the development of vocabulary acquisition and meaningful interaction in an L2.

Table 1. Profile of IVAs studies in foreign language education

Author	Type of Study	Sample & Context	IVA Type	L2 Aspect	Pedagogical Affordances	Challenges
Tai (2022)	Mixed methods study	89 EFL Chinese college freshmen	Google Assistant	L2 speaking	increased L2 oral proficiency in terms on fluency, decreased anxiety in L2 speaking tasks for low-level L2 learners mobile, ubiquitous L2 learning support	simple, uninteresting dialogue content/tasks Technological issues: recognition errors, irrelevant responses, distraction
Hsu et al. (2021)	Experimental design	50 Taiwanese EFL college learners	Amazon Alexa	L2 speaking & listening	enhanced L2 speaking proficiency, positive L2 learners' perceptions	IVAs' mispronunciations, provision of late responses
Gonulal (2021)	Quantitative study	42 Turkish university learners	Google Assistant	L2 learners' perceptions of IVAs	IVAs as highly useful agents to improve both their overall English and L2 pronunciation and speaking skills& vocabulary knowledge	limitation in dialogue-enhancing features such as humor capacity
Dizon (2020)	Experimental design	37 Japanese EFL 1st & 2nd university learners	Amazon Alexa	L2 speaking & listening	enhanced L2 speaking proficiency, ease of use & useful for learning	IVAs rate of speech, level of vocabulary & inability for modified output
Dizon & Tang (2020)	Mixed-methods case study	14 Japanese EFL university learners	Amazon Alexa	L2 learners' perceptions of IVAs in the EFL classroom	autonomous L2 learning in terms of skills, interaction, vocabulary enjoyment, information usability, effectiveness, satisfaction	Technical issues Understanding Reliability (being understood)

Moussali & Cardosso (2020)	Mixed methods study	11 EFL university learners	Amazon Alexa	L2 speaking	comprehensible and intelligible communication in L2 English	communication breakdowns due to pronunciation issues
Chen et al. (2020)	Mixed methods study	29 EFL college learners	Google Assistant	L2 learners' perceptions of IVAs	enhance listening & speaking skills, pronunciation, vocabulary more exposure to authentic English fun, easy to use, convenient, less stressful L2 tool	mechanical responses, could not imitate human–human interaction pronunciation errors hindered communication no corrective feedback no visual aids
Tai & Chen (2020)	Mixed methods study	112 EFL 8th grade high-school learners	Google Assistant	L2 learners' willingness to communicate	improved WTC, low levels of anxiety high communicative confidence fun, accessibility, practicality	minor technical problems (e.g. internet speed, ASR system) limited linguistic ability
Underwood (2017)	Qualitative study	11 EFL primary learners	Amazon Alexa Siri Google Assistant	L2 speaking	meaningful and joyful interaction in L2 English	fast speech rate and inaccurate voice recognition
Moussali & Cardosso (2016)	Mixed methods study	4 EFL learners	Amazon Alexa	L2 learners' perceptions of IVAs	fun, easy to set up and use, creates authentic contexts, options for self-regulation, motivation, opportunities for input and output practice	problems with the speech recognizer

Hsu et al.'s [30] experimental study showed the impact of Amazon Alexa Echo Show on listening and speaking skills along with learners' perceptions towards Alexa's use for language learning purposes on 50 L2 Taiwanese college learners. The experimental group received seven Alexa sessions while all participants were asked to take pre-and post-mock TOEIC listening and speaking tests and complete a survey questionnaire. Results replicate Dizon's [16] findings in the experimental group, demonstrating a significant effect of IVA use on L2 learners' speaking but not on their listening ability which was mainly attributed to the opportunities provided by learner-IVA meaningful interactions to intermediate-level L2 learners to brainstorm on meaning, receive feedback, notice errors and modify their language. In relation to L2 learners' perceptions of using IVAs in the EFL class, findings concur with Dizon [16], Dizon and Tang [15], and Underwood [70], with learners highlighting Alexa's usefulness in the development of L2 speaking skills for specific purposes (e.g., presentation skills). The same positive disposition is also reported in Moussali and Cardoso's [52] small-scale study where Amazon Echo was considered a user-friendly, enjoyable, and helpful pedagogical tool for language learning, providing opportunities for input exposure and output practice while motivating learners to learn on their own. This is also highlighted in Moussali and Cardoso's [53] follow-up study that assessed Amazon's Alexa ability to recognize and process the different accents of non-native accented speech based on the accuracy of IVA's answers for pre-set questions. Results indicated that L2 learners with differing levels of linguistic proficiency faced overall no significant intelligibility issues in their communication with Alexa, as it could easily understand and accommodate accented speech, effectively detect pronunciation and lexical issues, and promptly provide learners with implicit feedback, prompting them to detect erroneous forms in their production of the target language.

Chen et al. [11] showed the effect of language proficiency on 29 L2 Taiwanese college learners' perceptions toward using Google Assistant (GA) when interacting with it for EFL learning purposes. In line with the findings reported by Dizon [16], Moussalli & Cardoso [52], [53], and Underwood [70], analysis of the data revealed L2 learners' overall favorable viewpoints toward the use of GA, which enticed their interest in considering ways to develop their vocabulary and oral skills. The perceived pedagogical utilization of GA was heavily influenced by the degree of mutual comprehensibility that L2 learners with different language proficiency levels achieved with GA. Contrary to Moussalli and Cardoso [53], results demonstrated that learners with higher language proficiency tended to benefit more from their GA interactions as they considered themselves better understood by the IVA than low-level learners who faced more challenges largely due to mispronouncing particular words. As in Hsu et al. [30], interacting with GA was found to be useful for intermediate and upper-intermediate L2 learners enabling them to identify their pronunciation errors or mistakes and offering more exposure to authentic pieces of conversational exchanges. The same finding was yielded in Gonulal's [23] study where intermediate to upper-intermediate EFL Turkish learners also seemed to have benefitted immensely from their interactions with an IVA in L2 learning in terms of L2 oral fluency and vocabulary acquisition.

GA has also been targeted in two related studies and studied for its impact on EFL learners' willingness to communicate (WTC) [65] and learners' oral proficiency outside the classroom [64]. In line with [53], high-school L2 learners in Taiwan in the

former study with low WTC showed more willingness to interact in English during Google-Assistant-language-learning (GALL) activities, were more confident to interact with other learners in English, and asked for help, stating that their interaction within the less threatening environment provided by GA helped them develop L2 fluency by lowering their levels of speaking anxiety. The effect of GA built-in on the smartphones of 89 Chinese college EFL learners' out-of-class oral proficiency was also studied by [64] via self-directed interactive activities over a period of one semester. Findings indicated that out-of-class use of GA significantly enhanced EFL learners' oral proficiency in terms of (i) *fluency* promoted within GA's anxiety-free and interactive environment in parallel to [53] and [65]; (ii) *content and vocabulary* that was supplemented by L1 and L2 learning support and multimodal feedback provided via audio, text, and visual aids; (c) *pronunciation*, due to the opportunities provided by GA for L2 learners to practice speaking accompanied by multimodal presentation of feedback akin to [16] and [52], [53] findings; and (d) use of simple grammatical structures. No significant improvement was found for high-proficient EFL learners' L2 speaking skills due to the simplistic, non-challenging nature of dialogue content used for practice between learners and the IVA.

4.3 Barriers in IVas implementation in EFL

Common barriers that hinder the effective integration of IVAs in the L2 learning process can be classified into three distinct categories based on the relevant reviewed studies:

(a) *communication breakdowns in IVA-learner interactions*. L2 learners most often reported difficulties regarding the mutual comprehensibility issues that hindered communication with the IVA, thus engendering feelings of distraction from the learning process [64] and leading to abandonment [15]. Technological issues such as fast rate of speech [70], [16], the advanced level of vocabulary [16], pronunciation errors [54], [11], mispronunciations and late responses [30], and inaccurate voice recognition [70], [52], [64] that often result to inappropriate search results when learners speak simultaneously were identified as the most common causes for miscommunication in learner-IVA interactions in EFL settings. Such issues were primarily found to be associated with pronunciation errors, pauses in speech, wrong sentence structure, and stuttering preventing the virtual assistant from understanding learners who determined to overcome these difficulties resort to repetition, rephrase a command or pronounce words differently [54], [11].

(b) *IVAs inability to imitate human-human interaction*. Although the utility of IVAs is presented as a useful, easy-to-use, and convenient AI tool to promote speaking-listening skills in the L2 context, its limited linguistic abilities [65] in terms of simplistic modified output as evidenced in the production of IVAs' mechanical [11] and often irrelevant responses [65], negatively impacts L2 learners' opportunity to practice extensively their L2 speaking skills on an individual, on a less threatening basis given the absence of available corrective feedback to enhance their efforts. GA's limited capacity to provide unique humorous responses to users' utterances or commands has also been reported to significantly downgrade EFL

learners-IVA interactions leading to 'humor fatigue' over IVA's canned responses to joke requests [23].

(c) technical issues that were mainly related to Wi-Fi or connectivity issues [15], internet speed, and problems with the Automatic Speech Recognition Technology embedded in IVA systems [65].

4.4 Research trends in IVas in EFL

The mixed methods research design was predominantly employed in 6 out of the total 10 L2 studies on IVA reviews used here (Table 1) based on survey questionnaires and semi-structured interviews for data collection involving samples of L2 learners with low-level English language proficiency that ranged between 11 and 122 L2 learners. The experimental method was also used in two studies [16], [30] involving the administration of pre-and post-L2 speaking and listening tests to examine differences in performance between experimental and control groups before and after the implementation of IVA intervention in the L2 learning process. Finally, small-case studies were undertaken in two instances involving a small number of participants deploying either qualitative [70] or mixed data-gathering approaches [52].

5 Discussion

This review identifies the pedagogical benefits and challenges associated with the introduction of IVA technology in FL learning and teaching practice and maps out current research trends in the field. With respect to our first research question, using IVAs was found to engage L2 learners in meaningful and joyful L2 English language interactions and improve their L2 listening skills both inside and outside the classroom, thereby incrementally leading to increased autonomous learning. Learners were favorably inclined toward the use of IVA for self-directed FL learning as IVAs were considered to be entertaining and easy to set-up and use systems, providing realistic contexts for human–machine interactions, adapting to learners' language learning needs (e.g., pronunciation issues) in a less threatening environment with learners exhibiting greater levels of participation, enthusiasm, confidence, as well as willingness to take risks when engaging in L2 conversational exchanges. With respect to the second research question, challenges related to the effective integration of voice-based IVAs in L2 education are mainly linked to the quality of learner-IVA interactions [16], as this was reflected in the communication breakdowns attributed mainly to IVAs inadequacies in their embedded Automatic Speech Recognition technology as well as to their linguistic abilities that obviate the provision of rich modified output to promote meaningful interaction and, by extension L2 learning. Key findings for the third research question indicated that mixed methods research design was the most predominant methodology adopted by most studies in the field, followed by the experimental design that used small sample sizes of varied linguistic ability in both instructional and informal language learning contexts.

However, as uncovered by our analysis, IVA integration, in the L2 education field is still relatively under-explored, suggesting the need for an interdisciplinary orientation in future research that will focus on IVA instructional design features and the underlying pedagogy for their deployment in EFL contexts [63]. Such research could be directed toward the following areas:

- the development of new techniques adopted to enhance the design and encoding of IVAs' responses database to natural language inputs as well as the increasing use of automated strategies for the acquisition and construction of databases using advanced technologies (e.g. Neural Networks) as a way to ensure L2 learner-IVA meaningful and authentic interactions [3]. Humanoid intelligent agents or 'Holographic AIs' proclaiming future advances in Augmented Reality (AR) have most recently emerged as a possible alternative to disembodied vocal IVAs, with a potential pedagogical expediency that needs to be further substantiated [28].
- the investigation of the extent to which L2 learners' cognitive and affective characteristics and contextual factors can influence active engagement and long-term gains in L2 learning with the aid of IVAs.
- a multidisciplinary theory-driven exploration of L2 learner-IVA interaction as a reciprocal effective and meaningful communication process leading to significant learning gains based on research in the areas of learning theory, psychology, and instructional communication.
- more longitudinal IVA intervention studies to evaluate their long-term effects on L2 learning-related aspects in terms of technical feasibility and pedagogical expediency involving focus group discussions to assess whether the users' learning needs and expectations are met [72].

Due to the nature of the review, selection, and filtering process, the following limitations can be associated with this systematic review: (i) it is likely that pertinent empirical research was missed (e.g. book chapters, conference proceedings) despite our thorough literature search. However, limiting our attention to only high-quality publications was to warrant that all studies included in our study had been subjected to a rigorous peer review process, (ii) included reviewed studies that contained the term 'intelligent virtual agent' as a descriptor in their title, abstract, summary or keyword list and were mainly written in English, thus related research reported in other languages was excluded, (iii) manual searches of certain CALL international peer-reviewed journals may have led to the omission or mistaken rejection of relevant articles, and (iv) the total of only ten IVA empirical studies in L2 learning leading to a cautionary approach with the interpretation of their results in EFL practice.

6   Conclusion

The aim of this review was to summarize current L2 interventionist studies on general-purpose voice-based IVAs addressing their pedagogical affordances and barriers inhibiting their effective integration in L2 educational practice. Evidence underscores the added pedagogical value of IVA newly-emergent technology in the

field of EFL learning promoting authentic interaction in the target language through increases in motivation and perceived novelty. Yet, as education enters the fourth industrial revolution era, substantial research needs to be expended on voice-based IVA instructional design and implementation in FL contexts for the improvement of specific L2 skills paving the way for eXtended reality (XR) language learning with the introduction of AR-based embodied virtual assistants.

7   References

[1] B.A. Abu Shawar and E.S. Atwell, "Chatbots: Are they really useful?," *Journal of Language Technology and Computational Linguistics*, vol. 22, pp. 29–49, 2007. https://doi.org/10.21248/jlcl.22.2007.88

[2] E. Adamopoulou and E. Moussiades, "An overview of chatbot technology," in *IFIP International Federation for Information Processing AIAI 2020, IFIP AICT 584* I. Maglogiannis et al., Eds. Springer Nature, pp. 373–383, 2020. https://doi.org/10.1007/978-3-030-49186-4_31

[3] L. Alfaro, C. Rivera, J. Luna-Urquizo, E. Castaneda, J. Zuniga-Cueva, and M. Rivera-Chavez, "New trends in pedagogical agents in education," in *International Conference on Computational Science and Computational Intelligence*, IEEE, pp. 923–928, 2020. https://doi.org/10.1109/CSCI51800.2020.00172

[4] T. Baker, L. Smith, and N. Anissa "Educ-AI-tion rebooted? Exploring the future of artificial intelligence in schools and colleges," Feb., 25, 2019. [Online]. Available: https://www.nesta.org.uk/report/education-rebooted/ [Accessed: Aug. 5, 2022]

[5] E. Balci, "Overview of intelligent personal assistants," *Acta Infologica*, vol. 3, no. 1, pp. 22–33, 2019. https://doi.org/10.26650/acin.454522

[6] S. Bibauw, T. Francois, and P. Desmet, "Discussing with a computer to practice a foreign language: Research synthesis and conceptual framework of dialogue-based CALL," *Computer Assisted Language Learning*, vol. 32, no. 8, pp. 827–877, 2019. https://doi.org/10.1080/09588221.2018.1535508

[7] T. Bickmore and J. Cassell, "Social dialogue with embodied conversational agents," in *Advances in Natural, Multimodal Dialogue Systems*, J. van Kuppevel, L. Dybkjaer, and N. Bernsen, Eds. Kluwer Academic, pp. 1–32, 2005. https://doi.org/10.1007/1-4020-3933-6_2

[8] V. Braun and V. Clarke, "Using thematic analysis in psychology," *Qualitative Research in Psychology*, vol. 3, no. 2, pp. 77–101, 2006. https://doi.org/10.1191/1478088706qp063oa

[9] D. Burden and M. Savin-Baden, *Virtual Humans: Today and Tomorrow*. (1st ed.). Chapman and Hall/CRC Press, Boca Raton, FL, 2019. https://doi.org/10.1201/9781315151199

[10] J. Cassell, "Embodied conversational interface agents", *Communications of the ACM,* vol. 43, no. 4, pp. 70–78, 2000. https://doi.org/10.1145/332051.332075

[11] H.H.-J. Chen, C.T.-Y. Yang, and K. K.-W. Lai, "Investigating college EFL learners' perceptions toward the use of Google Assistant for foreign language learning," *Interactive Learning Environments*, pp. 1–16, 2020. https://doi.org/10.1080/10494820.2020.1833043

[12] K. Cherry, "Overview of emotional intelligence," Sept. 24, 2018 [Online]. Available: https://www.verywellmind.com/what-is-emotional-intelligence-2795423. [Accessed: Sept. 1, 2022]

[13] I.-C. Choi, "Efficacy of an ICALL tutoring system and process-oriented corrective feedback," *Computer Assisted Language Learning*, vol. 29, no. 2, pp. 334–364, 2016. https://doi.org/10.1080/09588221.2014.960941

[14] R.R. Divekar, J. Drozdal, S. Chabot, Y. Zhou, H. Su, Y. Chen, H. Zhu, J.A. Hendler, and J. Braasch, "Foreign language acquisition via artificial intelligence and extended reality: Design and evaluation," *Computer Assisted Language Learning*, vol. 35, no. 9, pp. 2332–2360, 2021. https://doi.org/10.1080/09588221.2021.1879162

[15] G. Dizon and D. Tang, "Intelligent personal assistants for autonomous second language learning," *The JALT CALL Journal* vol. 16, no. 2, pp. 117–130, 2020. https://doi.org/10.29140/jaltcall.v16n2.273

[16] G. Dizon, "Evaluating intelligent personal assistants for L2 listening and speaking development," *Language Learning & Technology*, vol. 24, no. 1, pp. 16–26, 2020. https://doi.org/10.29140/jaltcall.v16n2.273

[17] E. Eckel, "Apple's Siri: A cheat sheet," June 8, 2019 [Online]. Available: https://www.techrepublic.com/article/apples-siri-the-smart-persons-guide/ [Accessed: Sept. 3, 2020]

[18] K. Frue, "SWOT analysis of smart speakers", January 29, 2019. Available: https://pestleanalysis.com/swot-analysis-of-smart-speakers/ [Accessed: Sept. 3, 2020]

[19] M. Fuckner, J.-P. Barthes and E.E. Scalabrin, "Using a personal assistant for exploiting service interfaces," in CSCWD'14: IEEE 18th International Conference on Computer Supported Cooperative Work in Design, Hsinchu, Taiwan, May 21–23 2014, pp. 89–94. IEEE. https://doi.org/10.1109/CSCWD.2014.6846822

[20] L.K. Fryer, D. Coniam, R. Carpenter, and D. Lăpușneanu, "Bots for language learning now: Current and future directions," *Language Learning & Technology*, vol. 24, no. 2, pp. 8–22, 2020. https://doi.org/10125/44719

[21] R. Ghnemat, A. Shaout, and A.M. Al-Sowi, "Higher education transformation for artificial intelligence revolution: Transformation framework," *International Journal of Emerging Technologies in Learning (iJET)*, vol. 17, no. 19, pp. 224–241, 2022. https://doi.org/10.3991/ijet.v17i19.33309

[22] M. Gilles and E. Bevacqua, "A review of virtual assistants' characteristics: Recommendations for designing an optimal human–machine cooperation," *Journal of Computing and Information Science in Engineering*, vol. 22, no. 5, 2022. https://doi.org/10.1115/1.4053369

[23] T. Gonulal, "Investigating EFL learners' humorous interactions with an intelligent personal assistant," *Interactive Learning Environments*, pp. 1–14, 2021. https://doi.org/10.1080/10494820.2021.1974489

[24] B. Goodman, A. Soller, F. Linton, and R. Gaimari, "Encouraging student reflection and articulation using a learning companion," *International Journal of Artificial Intelligence in Education*, vol. 9, pp. 237–255, 1998.

[25] A.C. Graesser, X. Hu, and R. Sottilare, "Intelligent tutoring systems," in *International Handbook of the Learning Sciences*, F. Fisher, C.E. Hmelo-Silver, S.R. Goldman, and P. Reimann, Eds., New York: Taylor & Francis, pp. 246–255, 2018. https://doi.org/10.4324/9781315617572-24

[26] A.C. Graesser, M.W. Conley, and A. Olney, "Intelligent tutoring systems," in *APA Educational Psychology Handbook, Vol. 3. Application to Learning and Teaching,* K. R. Harris, S. Graham, T. Urdan, A. G. Bus, S. Major, and H. L. Swanson, Eds. American Psychological Association, pp. 451–473, 2012. https://doi.org/10.1037/13275-018

[27] A.C. Graesser and D. McNamara, "Self-regulated learning in learning environments with pedagogical agents that interact in natural language," *Educational Psychologist*, vol. 45, no. 4, pp. 234–244, 2010. https://doi.org/10.1080/00461520.2010.515933

[28] X. Huang, F. Wild, and D. Whitelock, "Design dimensions for holographic intelligent agents: A comparative analysis," presented at the 1st International Workshop on Multimodal Artificial Intelligence in Education, MAIED 2021, 14 Jun 2021, Online.

[29] W. Huang, K.-F. Hew, and L.K. Fryer, "Chatbots for language learning – Are they really useful? A systematic review of chatbot-supported language learning," *Journal of Computer-Assisted Learning*, vol. 38, pp. 237–257, 2021. https://doi.org/10.1111/jcal.12610

[30] H.-L. Hsu, H.H.-J. Chen, and A.G. Todd, "Investigating the impact of the Amazon Alexa on the development of L2 listening and speaking skills," *Interactive Learning Environments*, 2021. https://doi.org/10.1080/10494820.2021.2016864

[31] G.J. Hwang and C.-Y. Chang, "A review of opportunities and challenges of chatbots in education", *Interactive Learning Environments*, 2021. https://doi.org/10.1080/10494820.2021.1952615

[32] G.J. Hwang, H. Xie, B.W. Wah, and D. Gasevic, "Vision, challenges, roles and research issues of artificial intelligence in education," *Computers & Education: Artificial Intelligence*, vol. 1, 2020. https://doi.org/10.1016/j.caeai.2020.100001

[33] J. Huang, G. Shen, and X. Ren, "Connotation analysis and paradigm shift of teaching design under artificial intelligence technology," *International Journal of Emerging Technologies in Learning (iJET)*, vol. 16, no. 5, pp. 73–86, 2021. https://doi.org/10.3991/ijet.v16i05.20287

[34] S. Hussain, O.A. Sianaki, and N. Abaneh, "A survey on conversational Agents/Chatbots classification and design techniques", in *WAINA 2019, AISC 927*, L. Barolli. Eds., Springer Nature, pp. 946–956, 2019. https://doi.org/10.1007/978-3-030-15035-8_93

[35] J. Jeon, "Exploring AI chatbot affordances in the EFL classroom: Young learners' experiences and perspectives," *Computer Assisted Language Learning*, 2022. https://doi.org/10.1080/09588221.2021.2021241

[36] J. Jeon, "Chatbot-assisted dynamic assessment (CA-DA) for L2 vocabulary learning and diagnosis," *Computer Assisted Language Learning*, 2021. https://doi.org/10.1080/09588221.2021.1987272

[37] W.L. Johnson and J.C. Lester, "Face-to-face interaction with pedagogical agents, twenty years later," *International Journal of Artificial Intelligence in Education*, vol. 26, no. 1, pp. 25–36, 2016. https://doi.org/10.1007/s40593-015-0065-9

[38] A. Kaplan and M. Haenlein, "Siri, Siri, in my hand: Who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence," *Business Horizons,* vol. 62, pp. 15–25, 2019. https://doi.org/10.1016/j.bushor.2018.08.004

[39] A. Khanna, B. Pandey, K. Vashishta, K. Kalia, P. Bhale, and T. Das, "A study of today's A.I. through chatbots and rediscovery of machine intelligence," *International Journal of e-Service Science Technology*, vol. 8, pp. 277–284, 2015. https://doi.org/10.14257/ijunesst.2015.8.7.28

[40] Y. Kim, J. Thayne, and Q. Wei, "An embodied agent helps anxious students in mathematics learning," *Educational Technology Research and Development*, vol. 65, no. 1, pp. 219–235, 2017. https://doi.org/10.1007/s11423-016-9476-z

[41] B. Kitchenham, O.P. Brereton, D. Budgen, M. Turner, J. Bailey, and S. Linkman, "Systematic literature reviews in software engineering – A systematic literature review," *Information and Software Technology*, vol. 51, no. 1, pp. 7–15, 2009. https://doi.org/10.1016/j.infsof.2008.09.009

[42] B.A. Kitchenham, *Guidelines for Performing Systematic Literature Reviews in Software Engineering (version 2.3)*. EBSE Technical Report, Keele University and University of Durham, 2007.

[43] C.D. Kloos, C. Alario-Hoyos, P.J. Munoz-Merino, M.B. Ibanez, I. Estevez-Ayres, and C. Fernandez-Panadero, "Educational technology in the age of natural interfaces and deep learning," *IEEE R. Iberoamericana Tecnologias Aprendizaje*, vol. 15, pp. 26–33, 2020. https://doi.org/10.1109/RITA.2020.2979165

[44] S. Lee, H. Noh, J. Lee, K. Lee, G.G. Lee, S. Sagong, and M. Kim, "On the effectiveness of robot-assisted language learning," *ReCALL*, vol. 23, no. 1, pp. 25–58, 2011. https://doi.org/10.1017/S0958344010000273

[45] R. Luckin, W. Holmes, M. Griffiths, and L.B. Forcier, "Intelligence unleashed – An argument for AI in education," 2016. [Online] Available: http://discovery.ucl.ac.uk/1475756/ [Accessed: Sept 6, 2022]

[46] D. Luo, "Guide teaching system based on artificial intelligence," *International Journal of Emerging Technologies in Learning (iJET)*, vol. 13, no. 8, pp. 90–102, 2018. https://doi.org/10.3991/ijet.v13i08.9058

[47] L. Massai, P. Nesi, and G. Pantaleo, "PAVAL: A location-aware virtual personal assistant for retrieving geolocated points of interest and location-based services," *Engineering Applications of Artificial Intelligence*, vol. 77, no. 1, pp. 70–85, 2019. https://doi.org/10.1016/j.engappai.2018.09.013

[48] N. Matsuda, E. Yarzebinski, V. Keiser, R. Raizada, G.J. Stylianides, W.W. Cohen, and K.R. Koedinger, "Learning by teaching SimStudent – An initial classroom baseline study comparing with cognitive tutor," in *Proceedings of the international conference on artificial intelligence in education*, G. Biswas and S. Bull, Eds., Springer, pp. 213–221, 2011. https://doi.org/10.1007/978-3-642-21869-9_29

[49] D. Moher, A. Liberati, J. Tetzlaff, and D.G. Altman, "The PRISMA group. Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement," *PLoS Med,* vol. 6, no. 7, 2009. https://doi.org/10.1136/bmj.b2535

[50] A. Morin, "What is social intelligence?" April, 7, 2020 [Online]. Available: https://www.verywellmind.com/what-is-social-intelligence-4163839 [Accessed: Sept 16, 2022]

[51] E. Mousavinasab, N. Zarifsanaiey, R. Niakan, S. Kalhori, M. Rakhshan, L. Keikha, and M. Ghazi Saeedi, "Intelligent tutoring systems: A systematic review of characteristics, applications, and evaluation methods," *Interactive Learning Environments*, vol. 11, pp. 1–22, 2018. https://doi.org/10.1080/10494820.2018.1558257

[52] S. Moussalli and W. Cardoso, "Are commercial 'personal robots' ready for? Focus on second language speech," in *CALL Communities and Culture – Short Papers from EUROCALL 2016,* S. Papadima-Sophocleous, L. Bradley, and S. Thouesny, Eds. pp. 325–329, 2016. https://doi.org/10.14705/rpnet.2016.eurocall2016.583

[53] S. Moussalli and W. Cardoso, "Intelligent personal assistants: Can they understand and be understood by accented L2 learners?," *Computer Assisted Language Learning*, vol. 32, pp. 1–26, 2020. https://doi.org/10.1080/09588221.2019.1595664

[54] B. D. Nye, A.C. Graesser, and X. Hu, "AutoTutor in the cloud: A service-oriented paradigm for an interoperable natural-language ITS," *Journal of Advanced Distributed Learning Technology*, vol. 2, no. 6, pp. 35–48, 2014.

[55] C.W. Okonkwo and A. Ade-Ibijola, "Chatbots applications in education: A systematic review," *Computers and Education: Artificial Intelligence*, vol. 2, pp. 1–10, 2021. https://doi.org/10.1016/j.caeai.2021.100033

[56] G.M. Perez and S. Saffon Lopez, "Exploring the uncanny valley theory in the constructs of a virtual assistant personality," in *Intelligent Systems and Applications*, Bi, Y., R. Bhatia, & S. Kapoor, Eds., Vol. 1037, Springer International Publishing, pp. 1017–1033, 2019. https://doi.org/10.1007/978-3-030-29516-5_76

[57] S. Pesty and D. Duhaut, "Artificial companion: Building an impacting relation," in *IEEE International Conference on Robotics and Biomimetics*. Karon Beach, Thailand, Dec. 7–11 2011, IEEE, pp. 2902–2907. https://doi.org/10.1109/ROBIO.2011.6181746

[58] K. Punch, *Introduction to Social Research: Quantitative and Qualitative Approaches*, Sage, 1998.

[59] J. Rickel, S. Marsella, J. Gratch, R. Hill, D. Traum, and W. Swartout, "Toward a new generation of virtual humans for interactive experiences," *IEEE Intelligent Systems*, vol. 17, no. 4, pp. 32–38, 2002. https://doi.org/10.1109/MIS.2002.1024750

[60] S. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach* (4th Ed.), Pearson, 2020. https://doi.org/10.1093/oso/9780190905033.003.0012

[61] S. Singh and H. Beniwal, "A survey on near-human conversational agents," *Journal of King Saud University – Computer and Information Sciences*, 2021. https://doi.org/10.1016/j.jksuci.2021.10.013

[62] R. Sottilare, A.C. Graesser, X. Hu, and B. Goldberg, *"Design Recommendations for Intelligent Tutoring Systems: Instructional Management* (vol. 2)," Orlando: Army Research Laboratory, 2014.

[63] W. Swartout, R. Artstein, E. Forbell, S. Foutz, C. Lane, B. Lange, J. Morie, F. Albert Rizzo, and D. Traum, "Virtual humans for learning," *AI Magazine*, vol. 34, no. 4, pp. 13–30, 2013. https://doi.org/10.1609/aimag.v34i4.2487

[64] T.-Y. Tai, "Effects of intelligent personal assistants on EFL learners' oral proficiency outside the classroom," *Computer Assisted Language Learning*, 2022. https://doi.org/10.1080/09588221.2022.2075013

[65] T.-Y. Tai and H.H.-J. Chen, "The impact of Google Assistant on adolescent EFL learners' willingness to communicate," *Interactive Learning Environments*, 2020. https://doi.org/10.1080/10494820.2020.1841801

[66] J. Taoum, A. Raison, E. Bevacqua, and R. Querrec, "An Adaptive Tutor to Promote Learners' Skills Acquisition During Procedural Learning. Workshop eliciting Adaptive Sequences for Learning (WeASeL)," in ITS'18, Montréal, Canada, June 11–15, 2018, fhal-02089212f

[67] G. Terzopoulos and M. Satratzemi "Voice assistants and smart speakers in everyday life and in education," *Informatics in Education*, vol.19, no. 3, pp. 473–490, 2020. https://doi.org/10.15388/infedu.2020.21

[68] V. Timpe-Laughlin, T. Sydorenko, and P. Daurio, "Using spoken dialogue technology for L2 speaking practice: What do teachers think?," *Computer Assisted Language Learning*, 2020. https://doi.org/10.1080/09588221.2020.1774904

[69] D. Traum, "Models of culture for virtual human conversation," in *Universal Access in Human-Computer Interaction: Applications and Services*, *Vol. 5616,* C. Stephanidis (Ed.), Springer, Berlin, Heidelberg, pp. 434–440, 2009. https://doi.org/10.1007/978-3-642-02713-0_46

[70] J. Underwood, "Exploring AI language assistants with primary EFL students," in CALL *in a Climate of Change: Adapting to Turbulent Global Conditions – Short Papers from EUROCALL 2017*, K. Borthwick, L. Bradley, and S. Thouesny, Eds., pp. 317–321, 2017. https://doi.org/10.14705/rpnet.2017.eurocall2017.733

[71] S. Varma, "Siri vs Google Assistant vs Alexa: Which is the smartest virtual assistant in 2018?," June 14, 2018 [Online]. Available: https://gadgets.ndtv.com/apps/features/siri-vs-googleassistant-vs-alexa-which-is-the-smartest-virtual-assistant-in-2018-1867292 [Accessed: September 17, 2022]

[72] G. Veletsianos, R. Heller, S. Overmyer, and M. Procter, "Conversational agents in virtual worlds: Bridging disciplines," *British Journal of Educational Technology*, vol. 41, no. 1, pp. 123–140, 2010. https://doi.org/10.1111/j.1467-8535.2009.01027.x

[73] Y. Xiao and M. Watson, "Guidance on conducting a systematic literature review," *Journal of Planning Education and Research*, vol. 39, no. 1, pp. 93–112, 2019. https://doi.org/10.1177/0739456X17723971

[74] K. Zhang and A.B. Aslan, "AI technologies for education: Recent research and future directions," *Computers & Education: Artificial Intelligence*, vol. 2, pp. 1–11, 2021. https://doi.org/10.1016/j.caeai.2021.100025

8 Authors

Eirene Katsarou is a full-time ESP instructor at Democritus University, Thrace in Greece. She holds a BA in English Language & Literature (Aristotle University of Thessaloniki, Greece), an MA in Applied Linguistics (University of York, UK), and a PhD in Descriptive & Applied Linguistics (University of Essex, UK). Her research interests include: L2 vocabulary acquisition and language learner strategies, individual differences research, research methods in applied linguistics, computer-assisted language learning.

Fridolin Wild is professor at the Institute of Educational Technology of The Open University, also leading the Performance Augmentation Lab (PAL). He has been leading numerous EU, European Space Agency, and nationally funded research projects, including ARETE, LAAR, AR-for-EU, WEKIT, TCBL, ARPASS, Tellme, TELmap, cRunch, Stellar, Role, LTfLL, iCamp and Prolearn. He leads the Special Interest Group on Wearable-Enhanced Learning (SIG WELL) of the European Association of Technology Enhanced Learning (EATEL) and chairs the working group on Augmented Reality Learning Experience Models (ARLEM) of the IEEE Standards Association as well as the Natural Language Processing task view of the Comprehensive R Archive Network (CRAN).

Areti-Maria Sougari is an associate professor of Applied Linguistics in the Department of Theoretical and Applied Linguistics, School of English, Aristotle University of Thessaloniki. Some of her work appears in *TESOL Quarterly*, *Language and Education*, the *Journal of Applied Linguistics*, and other journals. Her research interests include teaching English to young learners, language teaching methodology, teacher education and development, and teaching English as a lingua franca.

Paraskevi Chatzipanagiotou is an Assistant Professor of Educational Management at European University Cyprus (Department of Education) and Director of Distance Education Unit. She has a PhD in Educational Management (Aristotle University of Thessaloniki), an MA in Teachers Education (Aristotle University of Thessaloniki), an MA in Distance and Open University (Hellas Open University), and a BA in Philosophy, Pedagogy, and Psychology (Aristotle University of Thessaloniki). She is involved in a lot of researched-based projects and participates as a trainer in further education programs for school staff, principals, and trainers. Her research interests are mainly focused on the fields of school management and leadership, innovation management, and digital transformation.