

Feiwen Xiao, Priscilla Zhao, Hanyue Sha, Dandan Yang and Mark Warschauer*

Conversational agents in language learning

<https://doi.org/10.1515/jccall-2022-0032>

Received November 25, 2022; accepted February 20, 2023; published online March 28, 2023

Abstract: Due to advances in technology, conversational agents are emerging as intelligent spoken dialogue systems that simulate natural conversation with human beings. A growing body of literature has investigated the potential of conversational agents in enhancing language learning across multiple contexts. In this paper, a broad scoping review examining the current literature on conversational agents and language learning was conducted. This review mapped APA PsycINFO, ERIC and ProQuest Dissertations & Theses databases, which yielded 23 papers for further analysis. Our examination of these papers suggests that there are three main ways in which conversational agents are used for language learning. This review discusses these three approaches and points to directions that require further research to fully exploit the potential of conversational agents in language learning.

Keywords: CALL; conversational agents; language learning; natural language processing (NLP); technology

1 Introduction

Language educators have long tried to exploit artificial intelligence–based dialogue systems to help students practice communication, typically in written form via chatbots (for a review, see Huang et al., 2022). However, until recent years, these efforts were hampered by the technical limitations of such systems and their inability to understand natural oral language. In the last few years, however, these limitations have been diminished by market advances in natural language processing (NLP) and artificial intelligence (AI), leading to the widespread use of voice-based dialogue systems called *conversational agents*.

***Corresponding author: Mark Warschauer**, University of California, Irvine, USA, E-mail: markw@uci.edu

Feiwen Xiao, Pennsylvania State University, State College, PA, USA, E-mail: ffx5014@psu.edu. <https://orcid.org/0000-0002-7284-9527>

Priscilla Zhao, Stanford University, Stanford, CA, USA, E-mail: puzhao@stanford.edu

Hanyue Sha, Columbia University, New York City, NY, USA, E-mail: hs3298@tc.columbia.edu

Dandan Yang, Boston College, Boston, MA, USA, E-mail: dandan.yang.2@bc.edu

Unsurprisingly, educators and researchers have become attracted by the possibility of using conversational agents for language teaching and learning—either as off-the-shelf products or through customised conversational agents that they programme using widely available software. In this paper, we review the scholarship on the use of conversational agents for language learning and point to possible research directions for the field.

Though conversational agents can potentially be developed and used for any language, we focus our review on their use for English, which is the most commonly taught second language. We first define what conversational agents are and explain how they work. We then introduce our literature review process and share what we found in terms of the main uses of conversational agents for second language learning. We conclude by suggesting the implications of this work to date for educators, researchers and developers.

2 What are conversational agents and how do they work?

Conversational agents (CAs) are spoken dialogue systems that simulate natural conversation with human beings. To accomplish this, CAs use artificial intelligence to construe the meaning or intent of spoken utterances and to generate an appropriate verbal response. The use of AI for natural language understanding distinguishes conversational agents from more basic types of dialogue systems that rely on simple algorithms (such as the presence or absence of particular words) to formulate responses.

Conversational agents work through a four-step process. First, they use AI to transcribe speech to text. They then use natural language understanding techniques to determine the underlying meaning or intent of the utterance (now converted to text). Next, they use AI or pre-programmed selection to generate a response to the utterance and, finally, convert the response to speech, either through a computer-generated voice or by playing a previously recorded utterance.

Conversational agents can be used in a variety of situations to automate and streamline communication and tasks. Some examples include customer service interactions, where a chatbot can assist with common enquiries and help customers navigate a website or product; scheduling and appointment scheduling, where a chatbot can help users schedule appointments or meetings; and e-commerce, where a chatbot can assist with product recommendations and online purchases. Other

potential uses include providing information and assistance in healthcare, education and financial services, as well as in personal assistant applications.

Conversational agents are available in a number of commercial products, including Siri, Alexa and Google Assistant, which can be accessed through smartphones, tablets, smart speakers or other digital devices. Those who wish to program their own conversational agents can use a number of publicly available natural language understanding tools, including Google's Dialogflow and Amazon Lex.

3 Methodology

For this review, we have chosen to employ a scoping review methodology, as the topic of our review, conversational agents in language learning, is a rapidly evolving area of research, with a wide range of studies being conducted.

A scoping review is a method of systematically identifying, mapping and summarising the available literature on a specific research topic (Munn et al., 2018). It is useful for identifying gaps in the literature and can help inform the design of future studies. It is suitable for reviewing a wide range of topics and can be used when the research on a topic is diverse, rapidly evolving or not well established. Scoping reviews are particularly useful when research on a topic is in its early stages, as they can help identify key areas for future research and provide an overview of the current state of knowledge on a topic. Therefore, a scoping review is suitable for this paper, which focused on reviewing the literature around an emergent topic in language learning.

To conduct a broad scoping review, we sought all scholarship—including published papers, dissertations, theses, reports and pre-prints—that focused on the use of conversational agents in language learning contexts. We searched APA psycINFO, ERIC and ProQuest Dissertations & Theses databases for articles written in English that included a combination of at least one term referring to conversational agent technology and at least one term referring to a language learning context (see Appendix I for the search keywords). Given that conversational agents are a new technology that has only emerged and been developed in recent years, we only included studies that were published after 2010, and the database search yielded a total of 569 articles. On top of the database search, we also screened the references of 11 review and synthesis articles on relevant topics. We conducted two rounds of screening where we first screened the title and abstract of each article and then read through each article to determine if it actually met our criteria of focussing on conversational agents in language learning contexts. From this screening, a total of 23 papers were selected for the coding and review.

To conduct a comprehensive review of the screened-in articles, we coded the following information for each article: citation, the country where the study was conducted, study design (e.g. mixed methods, case study, randomized controlled trial), the type of conversational agent used in the study (e.g. Google Assistant or Alexa), the setting of the study (e.g. in a university or at home), the age of the participants, the first language of the participants, whether the participants were English as a Foreign Language (EFL) or English as a Second Language (ESL) speakers, sample size of the study, how CA was used in language learning (general communication practice, task-based language learning or structured pre-programmed dialogue), intensity of the intervention, students' learning outcomes and implications. The descriptive information for each article can be found in Appendix II.

4 Uses of conversational agents and results to date

Our search suggests three main uses of conversational agents for language learning: general communication practice, task-based language learning and structured pre-programmed dialogue (see summary in Appendix III). We review each of these in turn.

4.1 General communication practice

One use of conversational agents is to facilitate general communication practice, where learners have an open conversation with agents in English. With built-in autonomous speech recognition (ASR) systems, agents are able to understand and respond to spontaneous speech input, thereby simulating interpersonal communication and serving as conversation companions for English learners.

A growing body of research has explored the potential of employing conversational agents to provide additional dialogue opportunities in second language (L2) classrooms. Moussalli and Cardoso (2020) explored the pedagogical use of conversational agents in providing extra L2 learning opportunities as part of classroom activities. In this usability study, the learners interacted with an Alexa-equipped Echo for 30 min with a set of preset questions and self-generated prompts. Then, they filled out a survey regarding their perceptions and attitudes towards the Amazon Echo and were subsequently interviewed. As indicated in the survey, all the participants enjoyed interacting with the Echo individually. Furthermore, they agreed that the Echo could function as a supplementary pedagogical tool to maximise

personalised learning opportunities. The findings suggest that such conversational agents could provide learners with valuable input exposure, output practice, pronunciation feedback and authentic English conversations in a stress-free environment.

Scholars from Japan conducted a series of studies investigating English learners' autonomous use of Alexa-based smart speakers among university students (Dizon, 2017; Dizon & Tang, 2019, 2020). For example, Dizon (2020) examined the effect of an intervention that integrated an Echo Dot into university-level English courses. Specifically, learners in the experimental group received a tutorial session on how to use the Echo Dot speaker and then received a list of possible commands that they could try out with the speaker. They were also encouraged to generate commands by themselves. The intervention, consisting of 12 min of human-agent interaction each week, lasted over 10 months. The results indicated that, compared with the control group, the experimental group demonstrated a greater gain in English speaking skills based on a speaking proficiency test developed by Payne and Whitney (2002). Nevertheless, the agent did not outperform the teacher in promoting students' listening proficiency.

Instead of examining the use of a single intelligent speaker, Obari and Lambacher (2019) incorporated Amazon Alexa and Google Home Mini together into a Business English training programme for Japanese university students. Forty-seven undergraduates were divided into two groups. Both groups used mobile learning applications and social media other than Amazon Alexa and Google Home Mini. Specifically, Group 1 used Google Home Mini daily to improve their listening and speaking with the following programs: Best Teacher, Travel English, Let's Play Around with English and BBC/CNN news. Group 2 used Alexa daily to improve their listening comprehension and vocabulary skills using the following programmes: Kikutan, English Quiz by Arc, Liberty English and Kindle. The post-test results based on the Test of English for International Communication (TOEIC) indicated that both groups achieved higher scores in terms of listening, speaking and total scores. In particular, Group 1 achieved a listening score almost double that of the pre-test, while Group 2 achieved higher gains in reading, which fulfilled the expectations of the initial learning goals. Based on a perception survey, students believed that speaking with a conversational agent benefitted their overall English proficiency.

On top of studies conducted in the formal learning setting, Dizon and Tang (2019, 2020) conducted two case studies that examined university language learners' in-home use of Alexa. In the first case study (Dizon & Tang, 2019), the learners received a tutorial and then initiated the self-directed use of intelligent speakers over four weeks. Their verbal communication was retrieved through the user history page to allow for further analysis of the interactions. A series of follow-up surveys and interviews indicated that the learners considered Alexa as a favourable tool for

English learning, as it provided meaningful L2 interactions and targeted pronunciation feedback and enhanced awareness of the learning gaps. However, these learners did not interact with the agent very frequently, and playing music was the most common interaction pattern (Dizon & Tang, 2019). In the second case study (Dizon & Tang, 2020), the same procedure was carried out, but the 14 participants were given two months to interact with Alexa. Likewise, the study highlighted that students perceived the agent to be fun, effective and useful for English learning. However, many of them did not demonstrate sustained use of the CA for language learning in an in-home environment.

The above studies examined the use of Alexa-equipped agents in learning English as a second language. Meanwhile, the ubiquity and mobility of Google Assistant gave rise to research studies that investigated its potential to reshape the L2 learning experience outside the classroom. Tai (2022) carried out an empirical study to understand the effectiveness of Google Assistant in improving L2 English learners' speaking skills by comparing the agent–human interaction with human–human interaction in L2 learning. The participants, which consisted of 89 Chinese college freshmen, were divided into three groups: (1) interacting with the Google Assistant (GA group); (2) interacting with native English speakers (L1 group) and (3) interacting with other L2 English speakers (L2 group). Specifically, the GA group interacted with Google Assistant for 10 min twice a week over the course of a semester, and they had to upload the transcript of their interactions with the agent. As for the L1 and L2 groups, the participants interacted with native English speakers and L2 English speakers, respectively, within the same timeframe. The three groups took an English oral proficiency test before and after the intervention. The post-test was followed by a survey and interviews to capture learners' perceptions and experience regarding their interactions with Google Assistant, native English speakers and L2 English speakers. The results indicated that the GA group and L1 group achieved significantly greater improvement in the oral proficiency test, and the GA and L1 group scored similarly in terms of all the subdimensions tested in the oral test: fluency, content, vocabulary, pronunciation and grammar. It was suggested that Google Assistant could function similarly as a native English speaker to help L2 learners improve their speaking skills. This study reinforced the potential of the intelligent speaker as an efficient conversation partner.

To sum up, most of the studies showed that learners generally embraced the use of CAs in practicing conversation in L2. They believed that the CA had the potential to act as a conversation companion and offer extra learning opportunities both within and outside the traditional L2 classroom. They described their learning experience with the CA as fun and motivating. Another benefit is that human–agent interaction alleviates speaking anxiety compared with talking to human beings. However, due to the short research periods in prior studies, longitudinal research examining how

students' attitudes towards L2 learning with CAs may change over time is needed. In addition, the existing research has explored the use of CAs within formal L2 learning environments. Yet, the ubiquity of CAs within households calls for more research on students' in-home autonomous L2 learning. A handful of studies on this topic have shown that students failed to achieve sustained use of CAs at home; further research on this topic is necessary.

4.2 Task-based language learning

Another way of using conversational agents in second language learning is incorporating the agent into tasks. To complete the tasks, learners need to communicate with conversational agents in L2. Such meaning-focused communication provides learners with authentic conversation opportunities and spontaneous feedback, which ultimately benefits L2 development (for an example, see Ellis, 2003).

Researchers have explored ways to integrate conversational agents into interactive tasks in the formal English learning environment. A research team in Chinese Taiwan conducted a series of research studies to examine the effectiveness of agent-assisted task-based language learning in promoting adolescent EFL learners' English proficiency in terms of speaking, listening and willingness to communicate (WTC) (Tai & Chen, 2020, 2022a, 2022b). The team employed Google Assistant (GA) and its associated speakers as the learning devices. They designed interactive tasks based on Google Assistant can-do lists (<https://assistant.google.com/learn/>). Additionally, they identified several interaction styles of the CA based on these tasks: (1) interviewer (e.g., "What's your Zodiac sign?"); (2) narrator (e.g., some facts about space); (3) facilitator (e.g., have a movie quiz); (4) interlocutor (e.g., ask to dress right for an outfit idea); and (5) entertainer (e.g., play a song on YouTube). Learners were informed of these interaction styles during a training session before officially interacting with the agent, thereby better navigating the conversation with the agent while completing the tasks.

In the first study, Tai and Chen (2022b) investigated the impact of GA with two different feedback presentation modes on English speaking proficiency among 88 EFL learners in Chinese Taiwan. The participants were randomly assigned to three groups: (1) interacting with Google Home Hub, which provided both audio and on-screen feedback (GA-Hub group); (2) interacting with Google Home Mini (GA-Mini group), which provided only audio feedback; and (3) interacting with teachers and peers in formal classroom settings. Before the intervention, the experimental groups underwent training on how to interact with the agent. Then, all three groups were assigned identical speaking activities that engaged learners to narrate, describe, ask questions, or express opinions. In each session, the

participants completed two interactive tasks. In each task, alongside verbal interaction with the agent, they completed a written worksheet. They completed the first task with their peers in a group moderated by the CA, while the second task required them to work with the agent one on one. The results indicated that both experimental groups outperformed the control group, as they received maximized L2 input. In addition, the GA-Hub group, which received both audio and on-screen feedback, exhibited better gains in speaking scores. This highlights the importance of multimodal feedback in agent-assisted language learning.

In the second study, Tai and Chen (2022a) examined the effectiveness of GA with two different media presentation modes of responses on English listening comprehension. They divided 92 adolescent English learners in Chinese Taiwan into three groups: 1) interacting with Google Home Hub with audiovisual on-screen responses (GA-Hub group), 2) interacting with Google Home Mini with the audio-only response (GA-Mini group) and 3) listening to CD players. There were 40 interactive gamified listening tasks for learners to complete. In line with the first study, the participants were assigned two games in each session. They completed a written worksheet in addition to the games. The experimental groups collaborated on the first game but worked individually on the second one, while the teacher led the whole session in the control group. A follow-up interview suggested that learners perceived gamified tasks mediated by the CA to be enjoyable. Furthermore, the statistical analysis showed that the two experimental groups scored higher in listening scores in the post-test, with the GA-Hub group exhibiting a much more significant difference than the GA-Mini group. In particular, the listening comprehension scores of the GA-Hub group improved enormously over time. These findings confirmed the advantage of the CA in resembling meaningful interaction in L2, thus promoting listening comprehension through speaking and meaning negotiation (Long, 2015). Meanwhile, interacting with the CA in different roles grants exposure to different types of input, which leads to better learning outcomes. Additionally, the outperformance of Google Hub validated that multimodal representations of responses helped bridge the missing information, strengthen the connection between form and meaning and enhance information retention.

Apart from English skills, Tai and Chen (2020) explored the impact of Google Assistant on learners' willingness to communicate in English, an essential dimension of L2 learning and communicative competence. Compared with the two previous studies, they recruited a larger sample size of 112 eighth-grade EFL learners from a junior high school in Chinese Taiwan who were native Chinese speakers (aged 15–16 years). The participants undertook the same intervention procedure as the participants in their other studies: receiving training, completing eight different agent-assisted tasks (i.e., playing games with the agent, open conversation, trying out music commands) and participating in an interview. The

findings suggested that such agent-mediated learning environments lowered learners' affective filter and speaking anxiety, boosted their motivation and increased their WTC. In addition, most participants loved playing interactive games with GA, but not all the tasks enhanced their WTC. They preferred tasks that allowed them to express their opinions and ask questions and those contextualised in topics that overlapped with their background knowledge.

Furthermore, Galvan-Romero (2022) conducted a pilot study to understand the role of CAs in ESL learners' WTC within the Hispanic and Latin American migrant community. During the intervention, 10 participants used Google Assistant on their mobile phones to complete tasks assigned by a teacher over 10 weeks. The tasks required learners to gather information to prepare for the upcoming class and to conduct conversations regarding daily life topics. Questionnaires were administered before and after the intervention to gain insights into learners' perceptions of their motivation, anxiety, communication strategies and self-perceived competence. According to the preliminary findings, students reported that they would like a training session to learn more about the features of the CA. It was also reported that some features of human–human interaction were missing from the agent (e.g. asking for repetition or speech adjustment).

Likewise, Hsu et al. (2021) studied the impact of the Amazon Echo on the development of L2 speaking and listening skills among 50 college students in Chinese Taiwan. Both the control group and experimental group received the same vocabulary instruction: text-reading or movie-watching, explicit vocabulary instruction and vocabulary practice. The experimental group received an additional session in which the participant interacted with Alexa with given tasks. They received training at the beginning of this session and then engaged in the CA-assisted tasks, including obtaining information from Alexa, commanding Alexa to perform specific functions and playing games. Students were then asked to fill out a worksheet based on the assigned task. The data analysis showed that doing the tasks with Alexa significantly improved the speaking scores alone, which resonates with Dizon's (2020) findings. The study also revealed that the agent could reduce speaking anxiety and offer rich oral interactions in L2.

Previous studies either focused on students of similar proficiency levels or overlooked students' prior proficiency levels. Chen et al. (2020) investigated the learning experiences of students of mixed proficiency levels with Google Assistant. The participants, who were 29 college students, were asked to participate in six tasks with Google Home Hub. The tasks required learners to ask GA various questions for information, command GA to play music and tell stories, play interactive games and conduct an open conversation with GA. A follow-up interview and survey were used to further understand learners' perceptions. In line with the studies reviewed above, the participants generally deemed that interacting with GA was enjoyable and less

stressful than speaking to human beings. They also perceived GA's utterances to be comprehensible and natural and believed in GA's potential to improve their English listening and speaking skills. However, their perceptions varied according to their proficiency levels. Higher-level learners were more likely to achieve mutual comprehensibility with the agent, whereas lower-level learners encountered more challenges due to mispronunciation. The less capable learners suggested that more visual aids on the touch panel would help identify the linguistic gap and lead to better communication (Tai & Chen, 2022b).

Rather than using an off-the-shelf CA available on the market, Jeon (2022) designed a CA using Google Dialogflow that targeted younger English language learners in the Republic of Korea. The agent was able to perform four types of responses: corrective feedback, prompts, fallback intent and evaluation. The voice interface can be accessed through tablets. The 36 primary schoolers participated in a 16-week EFL course that incorporated the CA, with three 40-min classes covered every week. Students familiarised themselves with the chatbots over the first two weeks and learnt different topics each week throughout the remaining 14 weeks. In each lesson, students followed the interaction configuration in the following sequence: whole group, small group, and individual. Specifically, the whole class reviewed the target language points and engaged in dialogue practice in the first session. In the second session of the week, the students were given information-gap tasks and completed them collaboratively with their peers. Then, in the third session, the students completed the tasks individually and filled out a worksheet as part of the task. They shared their learning experiences with the customised CA in a follow-up interview, which was leveraged along with user log data on the tablets for data analysis. Generally, in children's perception, there were three types of affordances the CA was able to provide: pedagogical affordance, technological affordance and social affordance. However, students' perception of the CA, their L2 level and technology competence affected their attitudes, thus further reshaping the proportion and power of these affordances for language learning. For example, students who perceived the CA to be pedagogically valuable were less likely to be discouraged by technical limitations and more likely to actively interact with the CA, thus generating more meaningful interaction. Additionally, most students preferred to interact with the CA one-on-one, as this interaction is free from peer pressure. It is suggested that teachers adapt the task to accommodate students of different proficiency levels and encourage both human-agent interaction and human-human interaction.

Targeting a younger age group, Underwood (2017) explored the perceptions of 11 primary school-aged EFL students regarding the current and future usage of AI speakers in the classroom. In this teacher-led design-based research conducted over nine months, students frequently used a variety of smart speakers, including Alexa, Apple's Siri and Google Assistant from the teacher's end to support classroom

activities. They all deemed this process meaningful and interesting. A co-design task was utilised to help the children communicate their ideas on what the agent looked like and how it could be used. They expected the agent to be a friend who could play with them. Unfortunately, the tasks the children wanted the agent to engage in with them were underdescribed.

Most research concentrated on ESL learners' perceptions of using a CA through interactive tasks, opening the black box of the L2 learning experience from a cognitive and affective perspective. In contrast, Wu et al. (2020) explored the potential of CAs as conversation partners in L2 learning from a usability perspective by comparing the use of Google Assistant for non-native and native English speakers. The participants were 25 university students, of whom 13 were native Chinese speakers and 12 were native English speakers. They had to complete six tasks on a smart speaker and smartphone. The task types included 1) playing music, 2) setting an alarm, 3) converting values, 4) asking for the time at a particular location, 5) controlling device volume and 6) requesting weather information. These tasks were delivered to the participants as pictograms, prompting the speakers to conduct real-world interactions with the agent. After completing the tasks, they participated in an interview regarding their user experience. The findings revealed that the L2 speakers attributed the communication breakdowns to their own poor pronunciation and linguistic knowledge. Additionally, they preferred the smartphone-based agent, as it provided visual feedback and helped identify communication breakdowns while granting them more time to reformulate the utterances. Interestingly, L2 speakers' demands for multimodal feedback were identified in previously mentioned studies (Hsu et al., 2021; Tai & Chen, 2022a, 2022b).

In addition to students' perceptions, teachers' perceptions played an integral role in the integration of CAs into task-based learning in the L2 classroom. Teachers' beliefs about CAs shaped the ways they were implemented in the classroom. The study of Timpe-Laughlin et al. (2020) stood out among the existing literature, which mostly concentrated on learners' perceptions. They used a survey and focus group interview to investigate teachers' perceptions of CA-based tasks and the methods of implementing them in their instruction. The 16 participants, who were ESL teachers, first filled out a demographic questionnaire and completed four intermediate-level CA-based tasks embedded on a website, followed by an interview. The four tasks were: (1) playing a guessing game; (2) ordering at a coffee shop; (3) making a request to a boss; and (4) disputing a billing error. Specifically, the first task required the learners to produce yes/no questions to find out a target character among eight characters preselected by the system. The CA would provide feedback in the form of affirmative and negative answers, and the interaction persisted until the user got the correct answer. In the second task, the learners interacted with the cashier in a coffee shop and ordered one food item and one drink item from the menu displayed on the

screen. In the third task, the learners had to call a boss to schedule a meeting and ask them to review presentation slides before the meeting. In the fourth task, learners were asked to call a service provider and dispute an error on their monthly bill. According to the results, all the teachers held positive attitudes towards the CA-based speaking tasks, and the third task turned out to be the most popular. In particular, they valued the authenticity of the critical feature of CA-based tasks and activities, where learners negotiated problems across multiple turns in the conversation. Furthermore, their perceptions were mediated by technology functionality, which directed their comments on whether the CA–human interaction was authentic or not. These teachers proposed two main ways of using the CA: (1) as supplementary practice materials outside of the classroom and (2) as potential diagnostics for formative assessments.

Overall, the majority of the studies utilised the tasks as a lens into learners' perceptions of L2 speaking and listening with CAs. Learners tended to perceive CAs as human-like and capable of communicating. In particular, the learners preferred one-on-one interaction with CAs over collaborative use in groups, as it created a stress-free learning environment and improved their learning motivation and willingness to speak in L2. Moreover, the learning process mediated by CAs was considered enjoyable, and learners tended to talk when CAs played the role of interlocutor in the conversation. Despite the promising nature of learners' perceptions, only one study targeted teachers' perceptions, the findings of which also reinforced the promise of CAs and uncovered their potential uses in task-based learning.

Despite the learning benefits of perception research, the existing literature has not systematically examined actual linguistic improvement through CA-based tasks. The task types were based on the off-the-shelf functions of commercial CAs. The most common tasks assigned to learners were to try out the commands, play games and ask the CA questions, which largely reflected the current available features of CAs. Only a handful of studies have redesigned the dialogue system to fit into specific learning contexts or redesigned the tasks for specific learning purposes. Nonetheless, the tasks emphasised the authenticity of the interaction and strived to promote meaningful output from the learners.

Among the most popular CAs available in the existing literature, L2 learners prefer CAs that provide both audiovisual and screen-based feedback. This multimodal design compensated for the language gap causing communication breakdowns.

4.3 Structured pre-programmed dialogue

A third way of using conversational agents in second language learning is through structured, pre-programmed dialogue. In this work, rather than using off-the-shelf

conversational agents, researchers or educators design and programme their own conversational agents to carry out dialogues on particular topics that are deemed relevant for desired learning outcomes.

Some of the most notable work in this area has been carried out by the Converse to Learn group at the University of California, Irvine and University of Michigan. This group has focused on broader goals related to young children's language and literacy development and content learning, but a major focus of their studies is the particular effects on second language learners. The group programmes its own tailored conversational agents using the publicly available Dialogflow from Google. In one series of studies, they deployed smart speakers and e-books that help young children, aged three to six years old, learn to read by asking them dialogic questions similar to what a parent or teacher might ask (Xu, Aubele et al., 2021; Xu, Wang et al., 2021). The original and follow-up questions follow a script designed to support children's learning of the story and vocabulary, with follow-up questions that are dependent on the conversational agents' capacity to interpret children's responses among a number of possible intents. The studies found that children in the AI-based dialogic condition learn much better than those without dialogue and approximately as much as those with a human asking the same questions, with the greatest benefit accruing to English learners.

In a second series of studies (Xu, Vigil et al., 2022), the team developed dialogic versions of a popular children's science television programme that children ages 4–6 watch on a tablet or laptop. The main character pauses periodically to ask the children questions, and then, using a similar design as above, continues the conversation based on their response. As above, the children in the dialogic condition learnt more content and vocabulary and were more positively affected than children who watched the traditional show, with English language learners once again demonstrating the greatest gains.

A number of scholars have replicated these types of structured dialogue studies with a more specific focus on second language learners. For example, Lee and Jeon (2022) carried out a study with 67 nine-year-old children studying English as a foreign language in the Republic of Korea. Using a similar approach to the UCI team, they programmed Dialogflow on a smart speaker to respond to a set of restricted questions from children, such as "What colour do you like?" or "Where is the hat?", and to ask the same question to the children, who would respond by referring to an illustrated worksheet. The conversational agent would carry on a pre-structured dialogue depending on whether the answer was correct, incorrect or not understandable. Then, to assess children's perceptions of the conversational agent and the activity, they were asked to draw and describe the conversational agent and to discuss their drawings and descriptions in an interview. The study found that more than 70% of the children perceived the conversational agent, although disembodied,

as having “human characteristics” and being capable of accurate second language communication. A number of these children, who had very little contact with native speakers of English, perceived that the agent could fulfil their need to have acquaintances who were fluent in English. As one stated, “I like the Google teacher because it would be good to have someone who is good at English around me”.

Finally, a team from Iowa State laid out pathways for designing and developing structured pre-programmed spoken dialogue systems for use in teaching (Chukharev-Hudilainen & Göktürk, 2020) and assessing (Chukharev-Hudilainen & Ockey, 2021) conversational ability among second language learners. For the purposes of instruction, Chukharev-Hudilainen and Göktürk suggested the following approach. First, the design team can create a sample instructional task and then record language learners performing the task. In their case, they recruited 9 adult non-native speakers of English to carry out a conversation about the availability of online learners. Then, they can perform linguistic analysis of this seed corpus to illuminate the kinds of conversational moves that typical conversation on the topic entails. Next, they can programme a conversational agent into a spoken dialogue system that performs similarly in response to the discussion comments, similar to how a human would respond through a set of recombinable templates. Then, they can test the system among themselves before field-testing it with end users.

Chukharev-Hudilainen and Ockey (2021) used a similar approach for developing a second language assessment system, which they field-tested with 42 adult English language learners. About half of the English learners preferred being assessed by a human, whose naturalness, body language, ease of exchanging ideas and better ability to understand their own imperfect English were appreciated. About half preferred being assessed by the conversational agent, whose slower, clear and standardised speech was appreciated.

This third area, structured pre-programmed dialogue, features the fewest papers, which is not surprising given the extensive preparation it involves. This is an expected growth area for future research and teaching.

5 Implications

The advancement of NLP techniques and AI algorithms grounded the development of conversational agents and transformed the dynamics of communication. Unsurprisingly, a rising number of researchers have begun to explore the significance and potential of using CAs in educational settings. Nonetheless, research in the L2 learning context is still emergent and scarce. This section further summarises the major findings from the existing literature and points to the implications for educators, researchers and developers.

5.1 For educators

Despite the ubiquity of conversational agents in daily life, it should not be assumed that all students have prior experience with conversational agents or have mastered the skills needed to interact with them. To ensure that everyone is on the same page, teachers are strongly encouraged to prepare a tutorial or training session for students to demonstrate interaction strategies and examples before students formally start dialoguing with CAs. If teachers choose to use a commercial agent, such as Google Assistant or Amazon Alexa, for instructional practices, they can refer to the current features and functionality of these commercial agents, which are available on their respective official websites.

Most of the research examined the use of conversational agents in the classroom setting, yielding promising results about learners' positive beliefs and perceptions of agent-based instructional activities. Nevertheless, most studies did not explore teachers' perceptions of CA-based tasks, thus calling for teachers' input on the implementation of conversational agents in future research. Despite being limited in size, current research on teachers' perceptions offers some insights insofar as the ways to use the conversational agent in instructional practices (Moussalli & Cardoso, 2020; Timpe-Laughlin et al., 2020). First, conversational agents can be used as potential diagnostic tools for formative assessment. Access to the transcript of human-agent dialogue from the device makes it possible for teachers to retrieve students' dialogues. Therefore, teachers can use it as a reference to evaluate students' oral proficiency and provide targeted feedback on their linguistic errors. Second, teachers can use the agent as a classroom assistant while motivating students to speak up, for example, by setting a timer, playing music, checking a fact or asking for definitions or spellings (Underwood, 2017).

Moreover, teachers can use CAs as a supplementary tool for individual practice, as they extend the L2 conversation opportunities outside the classroom. However, if learners use the conversational agent as a supplementary tool outside class, teachers should further consider how to help with sustained use of the CA. As implied by the research investigating students' self-directed use of the agent, students tended to lose interest in the conversational agent over time (Dizon & Tang, 2020). Managing students' outside-of-the-classroom learning experience would be challenging but meaningful. For example, teachers can monitor students' progress by asking them to do a daily check-in and collaborating with their parents, especially with younger learners' families. Teachers can even gamify the participation mechanics to get everyone engaged.

Moreover, given students' favourable reactions to the use of CAs in the classroom, educators should strive to integrate these agents seamlessly into classroom

tasks. The task types employed in recent research are very limited in variety and are mostly based on the available features of the commercial agents, with a focus on speaking skills. Teachers are encouraged to leverage the resources available to design tasks that force students to produce the target linguistic forms (Underwood, 2017). They should be sensitive to individual differences, such as learner proficiency levels and learning styles, when designing the tasks. According to existing research, younger learners prefer to work with the agent alone; learners of lower proficiency levels tend to participate less actively in the tasks (Jeon, 2022; Lee & Jeon, 2022). Therefore, teachers should provide personalised scaffolding and differentiated instruction in these agent-based tasks. Additionally, they should think about how to support multiple interactions in the classroom, including agent–human interaction and student–student interaction. Finally, most conversational agents do not target L2 learners and specific learning contexts. Teachers are encouraged to redesign the voice interface using the available tools to adapt to their teaching contexts.

5.2 For researchers

Notwithstanding the popularity and versatility of voice assistants, few research studies have explored the potential of conversational agents in the second language learning context. In the existing literature, the majority of the studies were case studies. The limited sample size and scope and the explanatory nature of the literature yielded mixed results. Regardless of this concern, the findings showed that a dominant proportion of learners felt less anxious and more motivated to learn with the agent. However, this might be dependent on the learners' characteristics and task types. Younger learners prefer to work with agents alone in information gap tasks, as it relieves them from peer pressure (Jeon, 2022). This finding is divergent from Underwood's (2017), where students spoke more when doing group work with the agent, though the task type is not specified. Adult college students who were involved in more structured tasks enjoyed working with peers because collaboration scaffolded their knowledge gaps (Tai & Chen, 2022b). Current findings highlight the importance of expanding the sample size, involving learners of diverse demographic backgrounds and proficiency levels and experimenting with more agent-based instructional practices across different learning contexts.

Furthermore, most studies examine students' L2 learning as mediated by the agent on a cognitive and affective level, rather than as tested via learning outcomes. A number of studies imply that L2 learners' speaking improved after the intervention of human–agent interaction, but the results were not validated by experimental evaluation. There were mixed findings about the effectiveness of conversational agents for listening skills (Dizon, 2020; Tai & Chen, 2022a). Several studies interested

in the role of conversational agents in students' willingness to communicate (Galvan-Romero, 2022; Tai & Chen, 2020) identified positive effects of the agent on this dimension of L2 learning. Future research calls for more experimental research with robust and systematic research methodologies. Apart from listening and speaking skills in general, researchers should pay more attention to other language skills (grammar, vocabulary, comprehension, etc.) and psychological factors that mediate the learning process (anxiety, motivation, self-confidence, etc.).

Most of the studies explored the use of conversational agents in the formal learning setting. Future research needs to shed light on the autonomous use of conversational agents outside the classroom and, if possible, conduct longitudinal studies on this topic. In this way, educators will be able to better locate where the intervention should come in to promote students' sustained and active use of conversational agents. Additionally, designers can become informed about what scaffolding should be implemented for self-directed use.

Moreover, to gain holistic insights into agent-mediated L2 learning, researchers need to leverage different perspectives rather than focus on learners' perceptions alone. Teachers' beliefs about technology shape their instructional practices and design. Yet, there exists a wide gap in this line of research, and only one study examines teachers' perceptions of agent-based tasks for L2 learning. When interacting with commercial conversational agents, the L2 learners also play the role of users. The findings of recent research indicate that user experience with the voice interface affects their perception along with the learning outcomes (Jeon, 2022). A possible research direction is to borrow perspectives from human–computer interaction to understand user experience.

5.3 For developers

Most commercial conversational agents, such as Amazon Alexa and Google Assistant, are designed for use by native speakers. However, non-native speakers have different interaction patterns than native speakers (Wu et al., 2020). This poses a challenge for the agent in achieving mutual comprehensibility with non-native speakers. According to recent literature, communication breakdowns persist as a barrier to successful communication, which shapes learners' perceptions and learning outcomes (Chen et al., 2020; Dizon & Tang, 2020; Tai & Chen, 2022a, 2022b; Wu et al., 2020). On one hand, there are unavoidable errors in non-native speakers' utterances related to syntax, pragmatics and pronunciation, among others; on the other hand, agents cannot recognise accented speech due to

the limited coverage of the language models in the current dialogue system. The mutual comprehensibility between the agent and non-native speakers is highly contingent on contextual factors, such as learners' proficiency level, age and the task type in which the agent is embedded (Chen et al., 2020). Developers should continue to extend language coverage in natural language processing systems and offer improved access to authentic interaction in English for non-native speakers. For pedagogical purposes, the agent should foster multi-turn dialogue instead of completing single-turn requests (Underwood, 2017). Ideally, a multilingual agent that can understand code-switching in both the learners' native language and their target language might better serve the learning purposes in a language learning setting (Jeon, 2022).

Given the unique characteristics of second language learners, emergent research has begun designing conversational agents targeted specifically at L2 language learners. To personalize the L2 language learning experience, agents should ideally have multiple versions of voice databases for speakers of different proficiency levels (similar to how, for example, there is a 'Simple English Wikipedia'). Multimodal representation of feedback can compensate for the anxiety caused by communication breakdowns, which could be especially useful for scaffolding low-proficiency learners (Jeon, 2022; Lee & Jeon, 2022; Tai & Chen, 2022a, 2022b). When learners struggle to make sense of the prompts from the agent, they can use on-screen modalities for meaning-making, thereby carrying the conversation forward. Moreover, learners prefer to interact with agents with human-like traits that provide affective feedback and reiterate their utterances (Xu & Warschauer, 2020). They also prefer interacting with agents that play the role of interlocutor in the conversation. Future studies might need to examine how different designs of embodiment and feedback modalities shape language learners' perceptions and learning outcomes. Overall, to better customise conversational agents to act as language partners for ESL learners, developers might need to collaborate with educators to co-design pedagogical conversational agents.

6 Conclusion

The study presented a comprehensive scoping review on the intersection of conversational agents and language learning. Through a rigorous screening process, we narrowed our focus to 23 papers, providing valuable insights into the three main types of usage of conversational agents: general communication practice, task-based

language learning and structured pre-programmed dialogue. The findings of the included studies indicate that conversational agents have significant potential as conversation companions, providing additional learning opportunities both in and out of the traditional L2 classroom. Furthermore, studies utilising tasks as a lens into learners' perceptions suggest that learners tend to perceive CAs as human-like and capable of communication. The studies reveal that the use of CAs through structured, pre-programmed dialogue shows promise as a method for improving L2 language learning.

We also summarise the key takeaways for educators, researchers and developers. For teachers, it is crucial to prepare students for how to interact with conversational agents before using them in the classroom. Studies suggest that conversational agents can be used as a diagnostic tool, classroom assistant and supplementary tool for individual practice. However, teachers should also consider the means of promoting sustained use and adapting the agent to their teaching context. Future researchers should focus on expanding sample size, involving diverse learners and experimenting with different agent-based instructional practices.

Finally, we note that the papers in our review were all published before the public release of ChatGPT, the large language model developed by OpenAI that has caught the world's attention. As we write this, developers and educators are beginning to experiment with creating voice interfaces with ChatGPT and other similar tools, such as GPT-3, and we can expect that this will open up a new avenue for the role of conversational agents in language learning.

In general, advances in natural language processing have created powerful new opportunities for the use of voice technology in second language learning. Artificial intelligence will never replace the need for human conversation but can augment human–human interactions in important ways. Our study acknowledges the current challenges of the development and implementation of CAs for L2 learners' English learning; however, we remain confident that this will be an important new direction to pursue for second language educators and researchers alike.

Research funding: This paper is based upon work supported by the National Foundation under Grants No. 1906321 and 2115382. The paper builds upon the last author's plenary address to the 17th International Conference on CALL in October 2021. We are grateful to ChinaCALL for the invitation to present this work at that conference.

Appendix I

Search Terms

Technology terms	Language learning terms
Conversational Agents	Language Learning
Automatic Speech Recognition	Language Acquisition
Intelligent Personal Assistants	Language Skills
Chatbots	English as a Second Language
Robots	English as a Foreign Language
Dialogue System	EFL
Voice Interface	ESL
Conversational System	Bilingual
Conversational Interface	Non-native Speaker
Natural Language Understanding Platform	Learn English
Voice Assistant	Language Learners
Smart Speaker	Computer-Assisted Language Learning
Conversational AI	
Siri	
Google Assistant	
Alexa	
Soapbox	
Dialogflow	
Amazon Lex	

Appendix II

Studies Included in this Systematic Review

Citation	Country	Study design	Type of CA	Setting	Age	First language	ESL/EFL	Sample size
Chen et al. (2020)	China	Mixed Method	Google Assistant	University	18–22	Chinese	EFL	29
Chukharev-Hudilainen and Göktürk (2020)	United States	Case Study	Custom Spoken Dialogue System	University	23–32	n/a	ESL	9
Chukharev-Hudilainen and Ockey (2021)	United States	Design Study	Custom Spoken Dialogue System	University	University Students	n/a	ESL	42
Dizon (2017)	Japan	Case Study	Alexa	University	University Students	Japanese	EFL	4
Dizon (2020)	Japan	RCT	Alexa	University	18–20	Japanese	EFL	31

(continued)

Citation	Country	Study design	Type of CA	Setting	Age	First language	ESL/EFL	Sample size
Dizon and Tang (2019)	Japan	Case Study	Alexa	Out of School	University Students	Japanese	EFL	2
Dizon and Tang (2020)	Japan	Case Study	Alexa	Out of School	University Students	Japanese	EFL	20
Galvan-Romero (2022)	UK	Pilot Study	Google Assistant	Out of School	n/a	Spanish	EFL	10
Hsu et al. (2021)	China	RCT	Alexa	University	College Students	Chinese	EFL	50
Jeon (2022)	The Republic of Korea	Mixed Method	PengTalk, SIA	Informal Setting	11–12	Korean	EFL	179
Lee and Jeon (2022)	The Republic of Korea	Qualitative Study	Custom CA programmed in Dialogflow	School	9	Korean	EFL	67
Moussalli and Cardoso (2020)	Canada	Case Study	Amazon Echo	n/a	n/a	n/a	EFL	4
Obari and Lambacher (2019)	Japan	Case Study	Google Home & Alexa	Business English Classroom	University Students	Japanese	EFL	47
Tai (2022)	China	RCT	Google Assistant	University	18–20	Chinese	EFL	1
Tai and Chen (2020)	China	Mixed Method	Google Assistant	English Class	8th Grade Students	Chinese	EFL	112
Tai and Chen (2022a)	China	RCT	Google Home	English Class	14–15	Chinese	EFL	92
Tai and Chen (2022b)	China	RCT	Google Home	English Class	14–15	Chinese	EFL	88
Timpe-Laughlin et al. (2020)	United States	Single Subject Design	Spoken dialogue system & CA	University	31–60	English	ESL	16
Underwood (2017)	Spain	Design-Based Inquiry	Alexa, Siri	Classroom	Primary School Students	Spanish	EFL	11
Wu et al. (2020)	Ireland	Case Study	Google Assistant	University	28.1	English & Chinese	EFL	32
Xu, Aubele et al. (2021)	United States	Randomized Block Design	Custom CA programmed in Dialogflow	Lab	3–6	English & Spanish	n/a	117

(continued)

Citation	Country	Study design	Type of CA	Setting	Age	First language	ESL/EFL	Sample size
Xu, Wang et al. (2021)	United States	RCT	Custom CA programmed in Dialogflow	Lab	3–6	English & Spanish	n/a	90
Xu, Vigil et al. (2022)	United States	RCT	Custom CA programmed in Dialogflow	Lab	4–6	English & Spanish	n/a	77

Appendix III

Three Main Uses of Conversational Agents for Language Learning

Theme	Studies included	Description	Principle research methods	Major findings
General Communication Practice	Dizon (2017) Dizon (2020) Dizon and Tang (2019) Dizon and Tang (2020) Moussalli and Cardoso (2020) Obari and Lambacher (2019) Tai (2022)	CAs used as a conversation partner to replicate human–human communication so that language learners can engage in natural verbal interaction.	Mixed-methods case studies	Learners enjoy interacting with these agents, which function as a supplementary pedagogical tool for personalised learning. CAs provide valuable opportunities for input exposure, output practice, pronunciation feedback and authentic English conversations in a stress-free environment. More impact seen in improvement of speaking skills than listening skills. Sustained use outside of the classroom is challenging to maintain.
Task-Based Language Learning	Chen et al. (2020) Galvan-Romero (2022) Hsu et al. (2021) Jeon (2022) Tai and Chen (2020) Tai and Chen (2022a)	Incorporation of CAs into interactive tasks to elicit meaning-focused interaction from language learners.	Randomised control trials Design-based inquiry Mixed-methods case studies	Studies focused on affect rather than learning outcomes, finding that students enjoyed these types of task-based interaction. Learners preferred CAs that provide both audio

(continued)

Theme	Studies included	Description	Principle research methods	Major findings
	Tai and Chen (2022b) Timpe-Laughlin et al. (2020) Underwood (2017) Wu et al. (2020)		Usability studies	and visual feedback. Tasks were selected based on publicly announced features of available CAs, with insufficient work on the development of tasks specific to the needs of learners.
Structured Pre-Programmed Dialogue	Chukharev-Hudilainen and Göktürk (2020) Chukharev-Hudilainen and Ockey (2021) Lee and Jeon (2022) Xu, Aubele et al. (2021) Xu, Wang et al. (2021) Xu, Vigil et al. (2022)	Use pre-programmed dialogue within media to achieve specific learning outcomes in comparison to the use of more traditional media.	Randomised control trials Case studies	Children interacting with the CAs are more engaged than those using traditional media. Children interacting with the CAs learn more language and content than those using traditional media, with second language learners demonstrating greatest gains. Children perceive CAs as hybrid agents combining human and machine elements.

References

Chen, H. H.-J., Yang, C. T.-Y., & Lai, K. K.-W. (2020). Investigating college EFL learners’ perceptions toward the use of Google assistant for foreign language learning. *Interactive Learning Environments*, 1–16. <https://doi.org/10.1080/10494820.2020.1833043>

Chukharev-Hudilainen, E., & Göktürk, N. (2020). Computers as conversational partners? Custom spoken dialog systems in applied linguistics. <https://doi.org/10.13140/RG.2.2.10017.68968/1>

Chukharev-Hudilainen, E., & Ockey, G. J. (2021). The development and evaluation of interactional competence elicitor for oral language assessments. *ETS Research Report Series*, 2021(1), 1–20.

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

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

Dizon, G., & Tang, D. (2019). A pilot study of Alexa for autonomous second language learning. *CALL and Complexity – Short Papers from EUROCALL 2019*, 107–112.

- 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.
- Ellis, R. (2003). *Task-based language learning and teaching*. Oxford University Press.
- Galvan-Romero, W. (2022). The role of intelligent personal assistants in migrant learners' willingness to communicate in English as a second language. In *CHI conference on human factors in computing systems extended abstracts, New Orleans, Louisiana, United States of America*.
- Hsu, H.-L., Chen, H. H.-J., & Todd, A. G. (2021). Investigating the impact of the Amazon Alexa on the development of L2 listening and speaking skills. *Interactive Learning Environments*, 1–14. <https://doi.org/10.1080/10494820.2021.2016864>
- 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.
- 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.
- 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*, 1–26. <https://doi.org/10.1080/09588221.2022.2067182>
- Long, M. (2015). *Second language acquisition and task-based language teaching*. Wiley-Blackwell.
- 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.
- Munn, Z., Peters, M. D., Stern, C., Tufanaru, C., McArthur, A., & Aromataris, E. (2018). Systematic review or scoping review? Guidance for authors when choosing between a systematic or scoping review approach. *BMC Medical Research Methodology*, 18(1), 1–7.
- Obari, H., & Lambacher, S. (2019). Improving the English skills of native Japanese using artificial intelligence in a blended learning program. *CALL and Complexity – Short Papers from EUROCALL 2019*, 327–333.
- Payne, J. S., & Whitney, P. J. (2002). Developing L2 oral proficiency through synchronous CMC: Output, working memory, and interlanguage development. *CALICO Journal*, 20(1), 7–32.
- Tai, T.-Y. (2022). Effects of intelligent personal assistants on EFL learners' oral proficiency outside the classroom. *Computer Assisted Language Learning*, 1–30. <https://doi.org/10.1080/09588221.2022.2075013>
- Tai, T.-Y., & Chen, H. H.-J. (2020). The impact of Google Assistant on adolescent EFL learners' willingness to communicate. *Interactive Learning Environments*, 1–18. <https://doi.org/10.1080/10494820.2020.1841801>
- Tai, T.-Y., & Chen, H. H.-J. (2022a). The impact of intelligent personal assistants on adolescent EFL learners' listening comprehension. *Computer Assisted Language Learning*, 1–28. <https://doi.org/10.1080/09588221.2022.2040536>
- Tai, T.-Y., & Chen, H. H.-J. (2022b). The impact of intelligent personal assistants on adolescent EFL learners' speaking proficiency. *Computer Assisted Language Learning*, 1–28. <https://doi.org/10.1080/09588221.2022.2070219>
- Timpe-Laughlin, V., Sydorenko, T., & Daurio, P. (2020). Using spoken dialogue technology for L2 speaking practice: What do teachers think? *Computer Assisted Language Learning*, 35(5–6), 1194–1217.
- Underwood, J. (2017). Exploring AI language assistants with primary EFL students. *CALL in a Climate of Change: Adapting to Turbulent Global Conditions – Short Papers from EUROCALL, 2017*, 317–321.
- Wu, Y., Rough, D., Bleakley, A., Edwards, J., Cooney, O., Doyle, P. R., Clark, L., & Cowan, B. R. (2020). See what I'm saying? Comparing intelligent personal assistant use for native and non-native language

- speakers. In *22nd international conference on human-computer interaction with mobile devices and services*, New York, NY, USA.
- Xu, Y., Aubele, J., Vigil, V., Bustamante, A. S., Kim, Y. S., & Warschauer, M. (2021). Dialogue with a conversational agent promotes children's story comprehension via enhancing engagement. *Child Development*, 93(2), e149–e167.
- Xu, Y., Vigil, V., Bustamante, A. S., & Warschauer, M. (2022). Contingent interaction with a television character promotes children's science learning and engagement. *Journal of Applied Developmental Psychology*, 81, 101439.
- Xu, Y., Wang, D., Collins, P., Lee, H., & Warschauer, M. (2021). Same benefits, different communication patterns: Comparing children's reading with a conversational agent vs. a human partner. *Computers & Education*, 161, 104059.
- Xu, Y., & Warschauer, M. (2020). Exploring young children's engagement in joint reading with a conversational agent. *Proceedings of the Interaction Design and Children Conference*, 216–228. <https://doi.org/10.1145/3392063.3394417>

Bionotes

Feiwen Xiao

Pennsylvania State University, State College, PA, USA

ffx5014@psu.edu

<https://orcid.org/0000-0002-7284-9527>

Feiwen Xiao is a first-year doctoral student at Penn State University. Prior to this, she received her master's degree in educational linguistics from the University of Pennsylvania. Her research deals with the intersection of educational technology, psychology, and linguistics. She is interested in leveraging artificial intelligence to promote language learning and STEM learning and empower students and educators in the digital learning age.

Priscilla Zhao

Stanford University, Stanford, CA, USA

puzhao@stanford.edu

Priscilla Zhao is a second-year Education Data Science Master student at Stanford Graduate School of Education. Her research focuses on the use of computer vision and artificial intelligence to promote positive parent-child interactions. Additionally, she is interested in using innovative technologies, such as large language models, to enhance STEM learning and improve the well-being of caregivers. Her ultimate goal is to empower young learners to reach their full potential and improve their lives through learning.

Hanyue Sha

Columbia University, New York City, NY, USA

hs3298@tc.columbia.edu

Hanyue Sha is a Master's Candidate at Teachers College, Columbia University, majoring in Bilingual/Bicultural Education. Hanyue's research interests include literacy and math development for at-risk bilingual children through various interventions.

Dandan Yang

Boston College, Boston, MA, USA

dandan.yang.2@bc.edu

Dandan Yang is a postdoctoral fellow at Boston College Lynch School of Education and Human Development. Dandan received her PhD from University of California Irvine with a specialization in language and literacy with technology. Her research focuses on developing and evaluations educational technologies for language learning and STEM education.

Mark Warschauer

University of California, Irvine, USA

markw@uci.edu

Mark Warschauer is a Professor of Education at the University of California, Irvine, where he directs the Digital Learning Lab. He received his doctorate in second language acquisition from the University of Hawaii. His research focuses on the use of digital media and artificial intelligence to promote language development, literacy, and learning among culturally and linguistically diverse learners. He is a member of the National Academy of Education.