

# A systematic review on artificial intelligence dialogue systems for enhancing English as foreign language students' interactional competence in the university

Chunpeng Zhai<sup>a</sup>, Santoso Wibowo<sup>b</sup>

<sup>a</sup> School of Engineering and Technology, Central Queensland University, Brisbane, Australia

<sup>b</sup> School of Engineering and Technology, CQUniversity Australia, 120 Spencer Street, Melbourne, VIC 3000, Australia

## ARTICLE INFO

### Keywords:

Artificial intelligence  
Dialogue system  
Foreign language learning  
Interactional competence  
University education

## ABSTRACT

Previous studies demonstrate that the use of artificial intelligence (AI) dialogue systems for English as a Foreign Language (EFL) education has effectively improved university students' reading, writing, and listening abilities. However, there are limited systematic reviews focused on the evidence-based interactional competence of EFL university students. This study aims to examine the use of AI dialogue systems to enhance EFL university students' interactional competence. Through the PRISMA process, this study identified 28 articles published between January 2013 and August 2022 in journals and conferences from the most popular databases, including Google Scholar, ProQuest, IEEE, ScienceDirect, and Web of Science. The systematic review identified six dimensions and 25 sub-dimensions that influence the application of AI dialogue systems for EFL learning. The six dimensions include technological integration, task designs, students' engagement, learning objectives, technological limitations, and the novelty effect. Gaps are identified that (1) components of debate and problem-solving skills in EF acquisition in university education seemed to be overlooked in the AI dialogue system design, and (2) the importance of embedding culture, humor and empathy functions were not taken into consideration in the AI dialogue system. This study finds that the development and implementation of an AI dialogue system in EFL is still in its infancy stage. Future research should emphasize meaning-based communication, intelligibility in language competency, debate, and problem-solving skills in university education.

## 1. Introduction

An artificial intelligence (AI) dialogue system is a software application that simulates natural human dialogue through the use of text or text-to-speech functions. This is accomplished with the assistance of natural language processing (NLP), which can be achieved through machine learning or deep learning (Zhai & Wibowo, 2022). Over the last decade, AI dialogue systems have been gathering momentum as studies show that these dialogue systems are easy to use, unbiased and in return boost users' confidence in a more friendly setting (Alsadoon, 2021; El Shazly, 2021; Hsu, Chen, & Todd, 2021). Nowadays, AI dialogue systems are used in a variety of domains, including health, marketing, business, retail, entertainment and foreign language learning (Hsu et al., 2021).

The development of novel language-learning applications can be traced back to the first dialogue system, ELIZA. ELIZA was developed by Joseph Weizenbaum at the MIT Artificial Intelligence Laboratory

between 1964 and 1966 as the first natural language processing computer software. Eliza was built to highlight the superficiality of human-machine interaction via the use of pattern matching to imitate conversation (Weizenbaum, 1976). The revolutionary ELIZA paved the way for other projects on chatbots, which are also known as text-based dialogue systems (Mastura, 2021). The introduction of AI Markup Language (AIML) has facilitated the design of dialogue systems for EFL learning. Some practitioners of computer-assisted language learning (CALL) can use established dialogue systems by appraising their language-learning potential based on the availability of an open-source framework (Mastura, 2021). Due to the wide developer community, other CALL practitioners are able to construct additional systems for language students, such as computer simulators in educational communication (CSIEC) (Mastura, 2021).

In the field of foreign language acquisition, several applications using AI dialogue systems have been developed to create interactive tasks for enhancing various aspects of a language learner's interactional

E-mail address: [c.zhai@cqu.edu.au](mailto:c.zhai@cqu.edu.au) (C. Zhai).

<https://doi.org/10.1016/j.caeai.2023.100134>

Received 12 September 2022; Received in revised form 18 March 2023; Accepted 18 March 2023

Available online 24 March 2023

2666-920X/© 2023 The Author. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

competence (Mastura, 2021; Timpe-Laughlin, Sydorenko, & Daurio, 2020; Young, 2011). Interactional competence is the capacity to use available language resources to deploy interactional resources as required by the speaker/hearer to convey their communication intents in real situational circumstances (Alsadoon, 2021). Interactional competence is the core of the communication abilities that EFL students are expected to acquire (Young, 2011). In his theory of interactional competence, Young (2011) conceptualizes interactional communication as including but going beyond the components of communicative competence. These components include: (a) rhetorical script; (b) the register of the practice; (c) the turn-taking strategies; (d) topic management; (e) the participation framework; and (f) means for signaling boundaries and transitions (Skogmyr & Balaman, 2018). In more ethnomethodologically oriented research within the field of interactional competence, researchers have focused on the observable, micro-level interactional aspects of speakers' competences (Girgin & Brandt, 2020; Hall, Hellermann, & Doehler, 2011; Skogmyr Marian & Balaman, 2018).

### 1.1. Interactional competence

Interactional competence entails understanding communication rules as well as utilization of both linguistic and interactional resources, such as turn-taking and task-related management, within a particular context (Galaczi & Taylor, 2018). In EFL spoken interactions, two interlocutors participate and co-construct in vigorous and reciprocal conversations (Galaczi & Taylor, 2018). Interactional competence reflects a type of interaction, such as collaborative and asymmetric, created by interlocutors in conversations by drawing from two perspectives: sociolinguistic-interactional and psycholinguistic-individualist perspectives (Roever & Kasper, 2018).

A psycholinguistic-individualist perspective focuses on an emphasis on the individual speaker and their communication skills, and it evaluates combinations of criteria such as fluency and pronunciation, lexical resources and grammatical variety, and topic management. However, the psycholinguistic-individualist perspective does not require that language usage be integrated in social interaction or situational settings (Roever & Kasper, 2018). This is a critically problematic for the real-world application of a language, since interlocutors do engage in a range of contexts and with a variety of interlocutors, and not merely to generate context-free fluent, correct, and complicated language.

A sociolinguistic-interactional perspective places an emphasis on the aims of speakers' linguistic choices as well as the impacts those choices have on their interlocutors (Roever & Kasper, 2018). It evaluates the appropriateness, conventionality, and efficacy of their utterances. It is reflective of sociolinguistic conceptions of language competency, with a concentration on the contextual and social dimensions of language usage. The sociolinguistic-interactional perspective is deemed to enhance interactional competence in real-world language usage that involves the ability to deploy language in interactive circumstances for social purposes and academic language (Roever & Kasper, 2018).

Researchers have applied the interactional competence theory to various language learning contexts, such as administering a paired oral discussion task within the context of language testing (Divekar et al., 2021, pp. 1–29; Huth & Betz, 2019; Kampittayakul, 2019). Through implanting these activities, the researchers were able to examine the test taker's knowledge and abilities situated in the sequentially of interaction and how learners can negotiate. The studies found that learning goals in the realm of interactional competence can be effectively taught and learnt in EFL class. In one study, Divekar et al. (2021, pp. 1–29) leveraged the interactional competence theory to engage students of foreign language learning in non-dyadic multimodal conversations within the cognitive immersive language learning environment. The findings demonstrate a statistically significant and sustained increase in vocabulary, understanding, and conversational abilities. Meanwhile, Kampittayakul (2019) investigated the use of translanguaging to

enhance interactional competence among Thai students of EFL in one-on-one tutorials. The findings suggest that 76 per cent of them improved their speaking skills, and 100 per cent of them identified convergence as the most prevalent function of listenership in order to demonstrate that they understood the material and to begin to converse.

However, existing studies found that university students in many EFL settings do not have many opportunities available to enhance their interactional competence outside of the classroom or even inside the classroom, as the spoken interactional practice is frequently limited to interactions between the teacher and classmates (Wahyuningsih & Afandi, 2020; Xie, 2017; Zrekat & Al-Sohbani, 2022).

An AI dialogue system for EFL is anticipated to play a significant role in influencing EFL learning at the university level in the near future (Roever & Kasper, 2018). Moreover, as the AI EFL dialogue system is commonly viewed as a tool that can leverage technical capacities to resolve difficult and complex foreign language learning problems, a growing body of studies is gaining traction on how AI dialogue systems for EFL can be used to support university students' interactional competence (Rosé & Ferschke, 2016). This development will affect how EFL students frame and examine issues, analyze multiple perspectives, and apply various resources via collaboration and interactions. Such learning environments require that EFL university students cultivate enhanced agency and cognitive, social-emotional, and behavioral skills for functioning in knowledge societies.

The main goal of EFL learning for students is arguably to acquire fundamental language skills to enhance the performance of communicative competence (Young, 2011). The term competence was first introduced by Chomsky (1965) in the field of linguistics in order to differentiate between knowledge of language in the abstract (competence) and the way in which knowledge is realized in the production and interpretation of actual utterances. However, Hymes (1972) challenged Chomsky's conception of competence, which he defined as knowledge of language apart from its use.

The concept put out by Hymes (1972) served as the foundation for an applied linguistic theory of communicative competence that was developed by Canale and Swain (1980). This theory tied language actions that took place in social circumstances to underlying knowledge. In applied linguistics, language testing, and language education, communicative competence was considered to be a feature of a particular person. This complex construct consisted of various component elements that separated one individual from others and made up communicative competence. Young (2011) proposed the interactional competence theory as an alternative theoretical framework to communicative competence.

The last decade has seen an unparalleled and exponential rise in the capabilities of technology and AI, which is posing a challenge to traditional university education (Moscardini, Strachan, & Vlasova, 2022). The use of an AI dialogue system for language learning is relatively new as students need to obtain trust and positive attitudes towards an EFL dialogue system (Bashori, van Hout, Strik, & Cucchiari, 2020). Engaging in interactional practices in the target language with dialogue systems can also reduce anxiety and increase the desire to participate in speaking practice (Bashori et al., 2020). There are prior studies on the use of dialogue systems to replicate a virtual target language-speaking environment. This was accomplished via two different kinds of learning activities: (a) a role-play activity and (b) a learning scenario representation (Huang, Hew, & Fryer, 2022; Wollny et al., 2021). However, there are several major issues relating to EFL learning for university students such as:

- (a) Lack of focus on improving EFL students' long-term learning objectives from spontaneous initiation of engagement with the dialogue systems,
- (b) Overlooked factors enhancing EFL students' wellbeing in the design of dialogue systems for EFL during academic life at university,

- (c) Lack of cross-cultural humor consideration in the dialogue systems for EFL,
- (d) Lack of cross-cultural awareness function in designing dialogue systems for EFL, and
- (e) Absence of cross-cultural empathy practice in designing dialogue systems for EFL.

How might universities, lecturers, researchers, training developers, and EFL dialogue system developers contribute to the advancement of EFL university students' interactional competence? Answering this question requires a more in-depth examination of how research studies with existing AI dialogue systems for EFL which are enhancing the interactional competence of EFL students and requires to conduct research studies with existing AI techniques. One important aspect to consider in evaluating AI dialogue systems for EFL is to examine the key dimensions that contribute to enhancing language learning experiences in universities. These dimensions represent essential features that need to be assessed and analyzed in order to provide a comprehensive and multifaceted understanding of the system's effectiveness. Some researchers have referred to these dimensions as "evaluation criteria" or "assessment indicators", which help guide the examination of AI dialogue systems in education (Adenle, Chan, Sun, & Chau, 2020; Da Silva, Fernandes, Limont, & Rauen, 2020; Gunnarsdóttir, Davidsdóttir, Worrell, & Sigurgeirsdóttir, 2020). In this study, the term "dimensions" refers to facets of developments, trends, or issues in EFL that need to be assessed and analyzed, encompassing technical, theoretical, and pedagogical perspectives. The development of clear and specific research questions can help to provide a more lucid understanding of these dimensions. For example, one research question might be:

What are the key dimensions of AI dialogue systems for enhancing student's EFL interactional competence?

To address this question, it is helpful to draw on the findings of previous studies that have employed multidimensional evaluation frameworks. For instance, Kuhail, Alturki, Alramlawi, and Alhejori (2023) utilized several dimensions, including educational field, platform, educational role, interaction style, design principles, empirical principles, and challenges and limitations, to evaluate an AI-based chatbot for English language learning. Similarly, Zhai and Wibowo (2022) proposed a framework for evaluating chatbots in language learning that includes dimensions such as embedding cultural, empathetic, and humorous dimensions in chatbot design, pedagogical effectiveness, and user satisfaction. Lastly, Li, Chang, and Wu (2020) used multiple dimensions, such as technical quality, pedagogical effectiveness, and learner satisfaction, to evaluate an AI-based English dialogue system. By synthesizing these findings, a comprehensive set of evaluation dimensions for AI dialogue systems in EFL can be identified. These evaluation dimensions questions might include:

- Pedagogical effectiveness: How well does the AI dialogue system support language learning outcomes?
- Interaction style: How do students interact with the AI dialogue system, and how does this affect their learning experience?
- Design principles: What principles guide the development of the AI dialogue system, and how do they contribute to its effectiveness?
- Cultural, empathetic, and humorous dimensions: How well does the AI dialogue system incorporate these elements, and what impact do they have on the learning experience?
- User satisfaction: To what extent are students satisfied with their experience using the AI dialogue system?
- Technical quality: How well does the AI dialogue system perform in terms of accuracy, responsiveness, and other technical aspects?
- Challenges and limitations: What challenges and limitations are associated with the use of the AI dialogue system, and how can they be addressed?

By focusing on these evaluation dimensions and developing research

questions around them, researchers can better understand the effectiveness of AI dialogue systems for EFL and contribute to the ongoing development of these systems in university settings. To address the existing gap and contribute to the comprehension of how AI dialogue systems have been utilized to enhance the interactional competence of EFL students, this systematic review analyzed studies that investigated interactional competence through the use of AI dialogue systems. This systematic review is guided by the research questions listed below:

1. What are the key evaluation dimensions of AI dialogue systems for enhancing student's EFL interactional competence?
2. What gaps arise from the results, and what additional study is required for achieving effective AI dialogue systems?

The growing global demand for English language learning and the limited opportunities for EFL students to practice their interactional competence in traditional settings highlight the need for an in-depth study of AI dialogue systems in university EFL contexts (Butler, 2019, pp. 477–496; Camiciottoli & Campoy-Cubillo, 2018; Kawinkoonlasate, 2020). This is because AI dialogue systems can offer personalized, accessible, and controlled learning environments that facilitate meaningful conversations and virtual immersion experiences (Tao et al., 2023). By focusing on interactions, researchers can better understand and optimize the potential benefits of AI dialogue systems, such as tailoring input and difficulty, providing real-time feedback, and enabling empirical research on EFL interactional practice (Divekar et al., 2021, pp. 1–29).

Furthermore, it is essential to conduct an in-depth study on the implementation of AI dialogue systems for EFL, particularly on the interaction strategies. This is important because interaction strategies can help to (a) promote spontaneous initiation of engagement and support EFL students' wellbeing during their academic life at university (Ulum, 2020), (b) embed cross-cultural humor detection and response functions in the design of dialogue systems for EFL (Zhai & Wibowo, 2022), (c) integrate cross-cultural awareness functions (Timpe-Laughlin & Dombi, 2020), and (d) introduce cross-cultural empathetic functions in EFL dialogue systems (Zhai & Wibowo, 2022). Zhai, Wibowo, and Cowling (2022), for example, explore the potential benefits, challenges, and methodologies associated with embedding humor and empathy into language learning systems. Wang et al. (2022) highlight the impact of a culturally-aware pedagogical agent on learners' emotions, perceived learning, and learning outcomes in computer-mediated communication. It emphasizes the importance of considering cultural aspects in the design of pedagogical agents to enhance the learning experience for EFL students. The study concludes that a culturally-aware pedagogical agent can positively affect learners' emotions, perceived learning, and learning outcomes. However, the importance of considering cultural aspects and other interaction components, such as humor and empathy, in AI dialogue systems for university EFL students has not been fully explored.

It is critical to investigate the role of the interaction components of humor, empathy, and culture in AI dialogue systems for university EFL learners. This is because these interaction components can help to: (a) facilitate learner motivation, (b) support intercultural competence and (c) address diverse learner needs. Huo (2022), for example, states that there is a positive impact of cross-cultural awareness functions on EFL learners' psychological well-being, academic involvement, and perseverance. A holistic approach to EFL dialogue systems, including cross-cultural humor recognition and empathic response features, contributes to an effective and engaging learning environment (Zhai & Wibowo, 2022). Zhang and Jing (2022) examine the use of AI technologies in facilitating cross-cultural communication and preserving intangible cultural heritage, demonstrating that AI technologies can bridge cultural gaps, translate languages, and foster an understanding of diverse traditions and cultural practices. Incorporating these aspects in EFL dialogue systems ensures that learners develop crucial intercultural

competence for effective communication in today's globalized world. Inkster, Sarda, and Subramanian (2018) highlight the potential of empathy-driven conversational AI agents like Wysa in promoting mental well-being, suggesting that such agents can serve as a valuable digital tool for supporting users in their learning journey. By considering emotional and cultural factors in the design of EFL dialogue systems, educators can better address the diverse needs of learners from various cultural backgrounds. Collectively, these studies have shown the importance of a holistic approach to EFL dialogue systems that considers a range of emotional and cultural factors for creating an effective and engaging learning environment.

This paper consists of six sections. Section two discusses the methodology of reviewing, selecting and data analysis of articles on dialogue systems. Section three presents an overview of the EFL dialogue system for interactional competence at the university level. This is followed by results of the findings on the interactional practice of dialogue systems in university education in Section four. Section five discusses the findings and gaps of the selected studies. Finally, section six entails theoretical and practical implications, limitations and recommendations for future studies.

## 1.2. Rationale of the chosen study

According to research that was released by The Brainy Insights, the global market for English language learning is predicted to increase from USD 11.35 billion in 2021 to USD 35.78 billion by 2030, at a compound annual growth rate (CAGR) of 13.6% throughout the period of forecasting that spans 2022–2030 (Insights, 2021). According to the British Council (Wright, 2022), there are 750 million people who speak English as a second or additional language. It is anticipated that the number of people studying English in different parts of the globe will continue to rise. According to the research titled "The English Effect" published by the British Council, there are presently 1.75 billion individuals worldwide who speak English language. This is one out of every four persons on the globe. It is predicted that by the year 2020, there will be 2 billion individuals utilizing the language. English as a Foreign Language (EFL) is becoming a working language for a growing number of influential people throughout the globe since it is necessary for the study of a wide variety of courses at universities. However, existing studies found that university students in many EFL settings do not have many opportunities available to enhance their interactional competence outside of the classroom or even inside the classroom, as the spoken interactional practice is frequently limited to interactions between the teacher and classmates (Wahyuningsih & Afandi, 2020; Xie, 2017; Zrekat & Al-Sohbani, 2022). Following concerns from stakeholders, the current teaching of spoken English and interactional competence is still viewed as unsatisfactory. This is due to the fact that many graduates still lack the ability to communicate in a foreign language, which prevents them from participating fully in transcultural and global society (Beshir & Yigzaw, 2022; Darmajanti, 2017).

University students in many EFL settings do not have many opportunities available to use the target language outside of the classroom or even inside the classroom, as the spoken interactive practice is frequently limited to interactions between the teacher and classmates. Xie (2017) used participant observation and questionnaires with rating and open-ended questions as research tools to study 128 EFL students from various academic disciplines taking business English classes. The findings indicate that approximately 24% of the university students reported that their limited English communicative skills prevented them from making full use of their target language with teachers and classmates in class. On the other hand, only 1.6% of the participants reported using English outside of the classroom between 80% and 100% of the time. According to the percentages, the amount of time spent by participants using their target language both inside and outside class is still insufficient. Wahyuningsih and Afandi (2020) found that universal students' progress in English, particularly in speaking, was hindered by

the fact that they seldom hear sounds and words spoken in English by people outside of the classroom. The authors believed that different social and cultural elements might result in a variety of input and communication sources for the first language. In this instance, the culture of students is characterized by a predominant use of the Javanese language in the students' communication within the communities. Because of this, students only get inputs in English sometimes outside of class. Zrekat and Al-Sohbani (2022) investigated interactional communication issues among 74 Saudi EFL students and revealed inadequate practice of English in and out of the classroom as one of the variables negatively impacting learners' speaking skills. The participants in this study (44.07 percent strongly agreed and 55.93 percent agreed) stated that there is a shortage of speaking practice both inside and outside of the classroom was identified as one of the variables negatively affecting learners' speaking abilities. In addition, the interaction is often hampered by large class sizes or by the use of the students' native language (Jafari & Ansari, 2012). Thus, some scholars suggest that AI technologies, such as dialogue systems, are required to facilitate EFL university students' expression and creativity (Bibauw, François, & Desmet, 2019; Jafari & Ansari, 2012).

Dialogue systems for EFL employ the interactional approach and provide a significant basis for EFL as conversation inherently provides chances for input, output, and interaction. The dialogue system delivers input and adjusts the degree of difficulty which may be tailored to the level of university EFL students. It is common that EFL students are required to articulate their intended meaning at each multi-dialogue turn which can be seen as a forced output (Bibauw et al., 2019). In classroom settings, Dialogue systems for EFL provide students the chance to practice meaningful conversations in a "virtual immersion" environment, which not necessarily as effective as having a conversation in the target language with a native speaker, but the opportunities can offer equivalent characteristics. University learners can spontaneously get access to the whenever they intent to, as the system does not lose its patience when faced with a speaker who is having difficulty (Bibauw et al., 2019). The dialogue systems also provide a learning environment that can be fully controlled and is potentially configurable towards optimal conditions on all impacting factors (feedback, learner modelling and adaptivity, and motivational support). This environment can be used not only for learning but also for research purposes. Dialogue systems for EFL have the potential to give completely monitored settings for conducting empirical research on EFL interactional practice as it eliminates the variable and unpredictable nature of a human interlocutor (Bibauw et al., 2019).

## 2. Systematic review method

This section conforms to the recommended items reported in the systematic review and provides insight into the applications of the dialogue system for EFL acquisition. A systematic review has been characterized as a form of knowledge synthesis that outlines exploratory research questions by methodically selecting, identifying, and synthesizing existing information (Colquhoun et al., 2014). A systematic review is useful for mapping the landscape of published literature and identifying research gaps, defining research objectives, and providing policymakers with suggestions (Tricco et al., 2018). This study aims to understand the advantages of a dialogue system for EFL learning to enhance long-term learning outcomes and promote university ESL students' well-being. To reach this goal, this study employs Wolfswinkel et al.'s (2013) five-step method for performing a systematic literature review and analysis. This five-step approach allows researchers to conduct a comprehensive search, assess and analyze the articles collected from databases. The five-step method entails (a) determining the scope of a review, (b) conducting a literature search, (c) choosing the final samples, (d) evaluating the samples using content analysis, and (e) presenting the study results.



## 2.1. Determining the scope of a review

This initial stage involves the clarification of the inclusion and exclusion criteria for relevant sources, as well as the criteria for identifying and retrieving relevant sources from the literature. To be included in the review, the journal articles needed to focus on the following selection criteria: (a) published in English in a full-text article, (b) associated with dialogue system with EFL in university, (c) focused on communicative or interactional competence (d) provided empirical data, and (e) published between January 2013 and August 2022. We selected the publication range between January 2013 and August 2022 due to the: (a) rapid development of AI and natural language processing (NLP), particularly with the emergence of deep learning and transformer-based models, (b) increasing adoption of AI dialogue systems for EFL education, and (c) evolving pedagogical approaches and global demand for English language learning which have prompted educators and researchers to explore innovative pedagogical approaches.

Articles are not considered when they (a) did not concentrate on dialogue systems for EFL in university, (b) focused on other language learning, (c) were written in languages other than English, (d) were editorials, (d) were opinion pieces, and (f) were dissertations. There are several reasons for excluding editorials, opinion pieces, and dissertations. First, the purpose of this systematic review is to provide a comprehensive, unbiased, and objective analysis of the available literature on a specific topic. [Firmstone \(2019\)](#) claims that systematic reviews need to focus on empirical research studies that follow a well-defined methodology and adhere to established research standards. Second, editorials and opinion pieces often represent the subjective perspectives and opinions of the authors ([Firmstone, 2019](#)). Third, [Siddaway, Wood, & Hedges \(2019\)](#) pointed out that editorials and opinion pieces generally lack the methodological rigor and detailed reporting of methods and results found in research articles, making it difficult to assess their quality and relevance to the review. This is supported by [Firmstone \(2019\)](#) who believes that the use of editorials, opinion pieces, and dissertations may introduce bias and potentially skew the overall findings of the review. In addition, [Thomas & Skinner, 2012](#) state that even though dissertations can contain valuable research, they are often not as extensively peer-reviewed as journal articles. The use of dissertations for conducting reviews may also increase the risk of including studies with methodological weaknesses or biases. [Table 1](#) summarizes and presents the inclusion and exclusion criteria.

The following criteria were used to select the articles from the databases for this study:

- (a) Does the paper discuss technologies used for dialogue systems aimed to enhance EFL students' performance of interactional competence during university education?

- (b) Does the paper discuss supporting learning activities used in dialogue systems aimed to improve EFL students' performance of interactional competence during university education?
- (c) Does the paper discuss EFL students' engagement with dialogue systems aimed to enhance EFL students' performance of interactional competence during university education?
- (d) Does the paper discuss learning outcomes of EFL students' engagement with dialogue systems focused the performance of interactional competence on during university education?
- (e) Does the paper discuss EFL students' perceptions and attitudes towards dialogue systems focused the performance of interactional competence on during university education?
- (f) Does the paper discuss challenges and problems of designing dialogue systems for EFL in relation to enhancing interactional competence during university education?

The reason for discussing the technologies used for dialogue systems, students' engagement, EFL students' perceptions and attitudes is because these technologies can play a significant role in increasing students' level of interactional competence, and subsequently improving their EFL students' performance during their university education ([Alsadoon, 2021](#); [El Shazly, 2021](#); [Hsu et al., 2021](#)).

## 2.2. Conducting a literature search

The second stage involves entails executing the search query against the chosen databases in order to get the search results. Several of databases including ProQuest, IEEE, Science Direct, and Web of Science were used for the review process. These databases were chosen based on their importance to education (e.g., Science Direct) and AI (e.g., IEEE) in order to assure the review's comprehensiveness. Google Scholar was selected to search the database simultaneously to expand our retrieval of data and Web of Science was included as a secondary database to identify additional material that may be relevant to our investigation.

An initial search was carried out by utilizing four sets of searches and each search terms is comprised of several variations of the term English as a foreign language as a medium of instruction, in which includes English as a foreign language (EFL), teaching English as foreign language (TEFL), teaching English to speakers of other languages (TESOL). The second set of search phrases focuses on concepts associated with interactional competence in which includes interactional skills, interactional tasks, interactional competence, conversational English, speaking skills. The third set of search phrases consists of terms associated with AI in which includes AI, natural language processing (NLP), neural network (NN), intelligent tutoring system (ITS), machine learning (ML), machine intelligence (MI), deep learning (DL), and classical algorithm output (CAO). The fourth set of search relates to the term of university education in which includes associate program, undergraduate program, academic program, diploma, further college program, postgraduate program, graduate diploma, master, and philosophical degree. However, when we combined these categories and put terms into the search bar on ProQuest, IEEE, Science Direct, Web of Science, and Google search engine, zero data was retrieved. Thus, we have only included search terms such as EFL, interactional competence, AI and university education. During the course of the search, a total of 211 publications were retrieved where they have an abstract or title that included all four sets of search terms.

There are several steps involved in the review process. Firstly, the authors restricted the publication period to the past ten years, from January 2013 to August 2022, to ensure that the literature is relevant ([Bener, 2021](#)). Then, duplicate copies of the research were removed, and this brought the total number of studies down to 180. Next, both peer-review journals and conference proceedings were included. In accordance with this criterion, the search did not include trade publications, editorials, books, or review articles. The search was further narrowed down to only include articles that were written in English

**Table 1**  
Inclusion and exclusion.

Inclusion	Exclusion
(a) Articles must be published in English in a full-text article,	(a) Articles did not concentrate on Dialogue systems for EFL in university,
(b) Articles must be associated with dialogue system and EFL in university,	(b) Articles focused on other language learning,
(c) Articles must focus on communicative or interactional practice,	(c) Articles were written in languages other than English,
(d) Articles must provide empirical data, and	(d) Articles were editorials on Dialogue systems for EFL in universities,
(e) Articles have to be published between January 2013 and August 2022.	(e) Articles were opinion pieces on Dialogue systems for EFL in universities, and
	(f) Articles were dissertations on Dialogue systems for EFL in university.

pertaining to EFL, and this brought the total number of results down to 115 papers. Following this, the titles and abstracts of the selected publications were analyzed to ensure whether or not they were pertinent to university students' interactional competence that made use of AI technology. According to this criterion, we excluded studies that involved non-EFL learning, e.g., focus on Turkish language learning by [Guzey, Yildiz, Demir, and Aksu-Dunya \(2022\)](#) non-communicative competence, e.g., examine the difference between media language learning and traditional language learning by [Tripathi, Ray, Sinha, and Ahmed \(2021\)](#), non-university level education, e.g., such as research carried out in senior high school by [Nirwana and Suhono \(2022\)](#), thesis writing, e.g., investigate learners' learning preference in class by [Arung, Rafi, and Dewanti \(2019\)](#), not student learners, e.g., focus on teachers' attitudes towards using computer assisted language learning by [Gorjian \(2017\)](#) or the use of non-AI technologies, e.g., the use of a blackboard program by [Al-Oqaily, Salam, and Na \(2022\)](#)). This resulted in 28 research being selected for the final round of consideration. [Table 2](#) illustrates the criteria for inclusion and exclusion along with the number of studies that were retained after determining whether or not the criteria were met.

### 2.3. Choosing the final samples

The third stage consists of selecting the final samples so that a full analysis may be performed on them. The search is confined to the titles and abstracts so that researchers can concentrate on the search results. All original article titles and abstracts are scrutinised for their relevance to AI dialogue systems which are used to enhance EFL university students' interactional competence. This resulted in the discovery of 211 related articles. Then, duplicate articles are eliminated. There are 28 articles remaining for further evaluation.

### 2.4. Evaluating the samples using content analysis

Based on the research objectives, a total of 28 articles were selected for further examination and evaluation. In this study, a systematic review was conducted in which data were gathered, processed, identified, and summarized. A six-step procedure was employed to determine the recurring dimensions. The first phase involves conducting a thematic analysis of the data to acquire a more in-depth understanding. Then, the initial codes were created. The third and fourth phases entail identifying sub-dimensions and reviewing the sub-dimensions. The fifth phase consists of compiling all of the pertinent concepts. Finally, the data were analyzed to ensure relevance to the study's objectives. To ensure relevance and review quality, the literature that emerged from the databases has been screened twice based on the PRISMA process. PRISMA stands for Preferred Reporting Items for Systematic Reviews and Meta-Analyses. It is a minimal set of criteria for reporting in systematic reviews and meta-analyses based on evidence. [Fig. 2](#) illustrates that the PRISMA statement contains a 27-item checklist and a four-phase diagram.

#### 2.4.1. Coding scheme

In order to emphasize on the interactional competence supported by AI dialogue systems for EFL at universal level, a mixed method combining both deductive or "top-down" and inductive or "bottom-up" approaches proposed by [Elo and Kyngäs \(2008\)](#) was adopted into our coding scheme. These methods have had significant application across many different fields of study, including human-interaction dialogue systems ([Diederich, Brendel, Morana, & Kolbe, 2022](#)), and the whole process of coding was carried out in Microsoft Word and Microsoft Excel. The three-step coding method proposed by ([Vollstedt & Rezat, 2019b](#)) was followed by the authors, who carefully read all of the transcripts and identifying dimensions of the retrieved articles until saturation was achieved. The three-step coding entails: initial coding, axial coding and selective coding are employed in identifying six dimensions supported

by AI dialogue system for EFL learning. Initial coding, often known as "open coding," refers to the initial phase in the coding process, especially for qualitative methodologies such as grounded theory. During this first phase, qualitative data is divided down into distinct fragments and assigned codes ([Vollstedt & Rezat, 2019b](#)). Open coding refers to the first interpretative procedure through which raw research data are classified and systematically evaluated. In grounded theory research, open coding was presented as a crucial methodological instrument for qualitative data analysis. ([Vollstedt & Rezat, 2019b](#)). Axial coding is a qualitative research approach that entails connecting data in order to identify codes, categories, and subcategories grounded in the voices of participants within gathered data. In other words, axial coding is one method for constructing data links ([Vollstedt & Rezat, 2019b](#)). Selective coding refers to the procedure of selecting one category as the core category and associating all other categories with it. The fundamental concept is to construct a singular plot around which everything else revolves. It is believed that such a fundamental notion always exists ([Vollstedt & Rezat, 2019b](#)).

Initial codes were generated through open coding, the authors refined the initial codes through iterative processes, including comparison, group discussion, and further code expansion. Along the process, initial coding schemes for AI dialogue systems for EFL were identified, such as technological integrations, task design, students' engagement, and learning objectives, as more articles were analyzed, they realized that the dialogue systems for EFL required further differentiation. These coding schemes allowed for a more comprehensive understanding of the various aspects and dimensions of EFL learning supported by AI dialogue systems ([Vollstedt & Rezat, 2019b](#)).

- (1) Technological Integrations: This code highlights the different technological components and tools that can be integrated into AI dialogue systems for EFL to enhance learning experiences, such as virtual reality, multimedia, and game-based learning environments.
- (2) Task Design: This code emphasizes the importance of designing engaging, meaningful, and contextually relevant tasks for EFL learners within AI dialogue systems to promote language development and authentic language use.
- (3) Students' Engagement: This code focuses on the level of interest, motivation, and involvement that AI dialogue systems can generate among EFL learners, which is crucial for successful language acquisition, and
- (4) Learning Objectives: This code addresses the specific goals and outcomes that AI dialogue systems for EFL aim to achieve, such as improving vocabulary, grammar, pronunciation, listening, speaking, reading, and writing skills.

As the authors analyzed more articles, they also encountered some potential limitations and disagreements in the literature. For instance, (1) Technological Limitations: This code represents the challenges and constraints faced by AI dialogue systems for EFL, such as imperfect natural language understanding, limited cultural awareness, or accessibility issues for some learners, and (2) Novelty Effect: This code captures the idea that the initial enthusiasm and engagement with AI dialogue systems for EFL may be driven by the novelty of the technology rather than its inherent educational value. Over time, this effect may diminish, causing a decrease in students' engagement and learning outcomes.

In the second stage of the research process, axial coding was employed by the authors as a systematic means of analyzing the findings derived from the initial open coding stage ([Vollstedt & Rezat, 2019a](#)). During this phase, each author independently examined the emergent codes, focusing on identifying relationships between them and organizing the data into meaningful categories and subcategories ([Harati, Nooshinfard, Isfandyari-Moghaddam, Babalhavaeji, & Hariri, 2019](#)). To ensure the rigor and reliability of the analysis, the authors engaged in

Table 2

General profile of the reviewed studies on AI dialogue systems.

No.	Authors and Year	Focus of Study	Research Design	Participants	Technological Integration	Task Design	Students' Engagement	Learning objectives	Technological Limitation	Novelty Effect
1.	<a href="#">Ayedoun et al. (2019)</a>	To foster students' readiness towards EFL verbal interactions	Survey	40 university undergraduate and graduates from Japan	DMS	<ul style="list-style-type: none"> <li>Dialogue interactions</li> <li>Negotiation for meaning</li> </ul>	Boost students' motivation for communicating in EFL	EFL students' engagement is enhanced towards communication	Unnatural sound Gibberish response	N/A
2.	<a href="#">Ayedoun, Hayashi, and Seta (2020)</a>	To enhance EFL students' wiliness to converse in the target language	Questionnaire	60 Japanese university students	DMS	<ul style="list-style-type: none"> <li>Negotiation for meaning</li> </ul>	Enhance students' participation in dialogue	The dialogue system motivates EFL students' willingness to communicate	N/A	N/A
3.	<a href="#">Chen, Yang, and Lai (2020)</a>	To investigate students' perceptions toward the use of IPA	Survey	29 university students from Taiwan	ASR & IPA	<ul style="list-style-type: none"> <li>Conversation</li> </ul>	Encourage students to speak with the dialogue system	A helpful tool to improve students' pronunciation and vocabulary	Difficult to understand heavy accented speech	N/A
4.	<a href="#">Divekar et al. (2021)</a>	To research immersive technologies and collaborative learning in a dialogue system	A mixed-method	10 university students from China	<ul style="list-style-type: none"> <li>ASR</li> <li>SDS</li> <li>CILLE</li> </ul>	<ul style="list-style-type: none"> <li>Cultural knowledge</li> <li>Role play</li> <li>Negotiation for meaning</li> </ul>	Enhance speaking and writing	Students' vocabulary retainment and conversation skills significantly improved Less anxious	Technology bugs	Yes
5.	<a href="#">Dizon (2017)</a>	To examine whether IPA, Alexa, is able to understand EFL English utterances	Interview Single group	4 s-year university students from Japan	<ul style="list-style-type: none"> <li>ASR</li> <li>Audiolingual</li> </ul>	<ul style="list-style-type: none"> <li>Storytelling</li> <li>Question &amp; answer</li> <li>Interview</li> </ul>	Students enjoyed talking with Alex	EFL students regards IPA as a useful tool to enhance language learning	Odd robotic voices	N/A
6.	<a href="#">Dizon (2020)</a>	To investigate speaking development, such as conversation, discussion, and presentation skills	Quasi-experimental design Single group	37 first- and second-year university EFL students from Japan	ASR	<ul style="list-style-type: none"> <li>Storytelling</li> <li>Question &amp; answer</li> <li>Negotiation for meaning</li> </ul>	Engage students to talk for more than 20 min	A useful tool to support listening comprehension and speaking proficiency	Unable to provide modifications	N/A
7.	<a href="#">Dizon and Tang (2019)</a>	Use a dialogue system, Alexa, for self-study	Qualitative method Single group	2 fourth-year university students from Japan	IPA	Conversation	Enjoy talking with Alex	Dialogue systems facilitate pronunciation	N/A	N/A
8.	<a href="#">El Shazly (2021)</a>	To examine the role of AI in interactional practices	Quasi-experimental Single group	48 students from Egypt	<ul style="list-style-type: none"> <li>ASR</li> <li>Audiolingual</li> </ul>	Role-play	Slightly enhanced students' speaking abilities	AI dialogue systems improved interaction and oral communication But increased instead reducing students' anxiety	N/A	Yes
9.	<a href="#">Hsu et al. (2021)</a>	To explore the effects of Amazon Echo Show on EFL students' listening and speaking skills	Quasi-experimental Single group	40 s or third year of university from Taiwan	IPA	<ul style="list-style-type: none"> <li>Read aloud</li> <li>Describe photos</li> <li>Answer questions</li> </ul>	Students tend to interact more with Echo	IPAs foster the students' speaking abilities due to highly interactive learning environment. Less stressful	Without visual displays	N/A
10.	<a href="#">Hsu, Chen and Yu (2021)</a>	To offer a chatbot system, TPBOT, for EFL students to reduce their fear of speaking English.	Pre- and post-experiment Control & experiment groups	100 students from Taiwan	Task-oriented chat robot system	Conversation	Increase students' engagement with TOBOT	The findings indicate that students are happy with this TPBOT and think that it has helped them to enhance their English-speaking abilities.	N/A	N/A
11.	<a href="#">Johnson (2019)</a>	To present two case studies in the data-driven development of Enskill®English for foreign language learning	Survey	209 university from Serbia and Croatia	D <sup>3</sup>	Conversation	Improve English proficiency to suit the demands of more advanced students	Enskill®English is useful for acquiring spoken English abilities.	Dialogue choices are too limited and restricted, and encountered bugs.	N/A
12.	<a href="#">Kim (2016)</a>	To develop an English-speaking system that	Survey	181 college students in Korea	Speech to text	Voice chat	Engage students in various voice chats	All of the participants showed statistically	N/A	N/A

(continued on next page)

Table 2 (continued)

No.	Authors and Year	Focus of Study	Research Design	Participants	Technological Integration	Task Design	Students' Engagement	Learning objectives	Technological Limitation	Novelty Effect
13.	Kim (2017)	improves students' speaking abilities To evaluate voice-based dialogue systems on EFL students' negotiation for meaning	Survey Between group	123 freshman university students from Korea	Audiolingual	<ul style="list-style-type: none"> <li>Questions and answer</li> <li>Negotiation for meaning</li> </ul>	Encourage students in meaningful dialogue	significant improvements in their speaking abilities. Dialogue systems play an important role in voice-based interaction using meaning negotiation approaches	Still use pre-determined topics and pre-stored phrases	N/A
14.	Kim, Cha, and Kim (2021)	To examine how AI chatbots affect students' speaking skills, motivate and shape students' speaking experiences	Pre- and post-experiment	49 university students	N/A technology integration. Only names Replika, Andy, Google Assistant	Asking and answering short questions or opinions	Motivate and guide students' speaking skills	The students' ability to speak was improved, especially their pronunciation, intonation, and stress.	N/A	N/A
15.	Li and Peng (2021)	To evaluate a blended teaching mode based on an AI learning platform	A mixed-method Single group	59 freshman university students from China	Audiolingual	Conversation	Positive attitudes can be found towards AI learning platform	Students EFL learning interests and attitudes towards AI dialogue systems are enhanced Anxiety was reduced.	N/A	N/A
16.	Li et al. (2020)	To examine the feasibility of replacing the conventional English-speaking method with the dialogue system	Survey	28 university students	TTS	Conversation	Continuous student engagement using AI dialogue systems	Positive perceptions of the system's user interface, learning style, and efficacy were recorded.	Some students had issues with their speech not being properly recognized.	N/A
17.	Mazzilli (2021)	To investigate how students perceive the chatbot Elbot and it affects the students' behavior in conflict-resolution skills	Questionnaire	28 third year university student learning German	CALL	Conversation	Improvement of conflict resolution abilities	The use of chatbots to enhance conflict-resolution skills in EFL	There were coherence issues during interaction with the chatbot.	N/A
18.	Mohammadzadeh and Sarkhosh (2018)	Through AI dialogue system to improve speaking ability	Quasi-experimental Between group	45 students from Turkey	<ul style="list-style-type: none"> <li>ASR</li> <li>SDS</li> </ul>	Conversation	Engage in problem-solving activities	Dialogue system enhances students' performance in speaking	N/A	N/A
19.	Morino et al. (2017)	To investigate the effect of digital storytelling on Japanese EFL students' attitudes and awareness toward CALL	Single group	120 university students from Japan	<ul style="list-style-type: none"> <li>ASR</li> <li>SDS</li> </ul>	<ul style="list-style-type: none"> <li>Storytelling</li> <li>Negotiation for meaning</li> </ul>	Enjoy storytelling conversations	The course increased students' awareness of CALL and improved their speaking proficiency.	Incapacity to have lengthy conversations	N/A
20.	Moussalli and Cardoso (2021)	To investigate API's ability to increase EFL students' phonetic awareness and allomorph production	Quasi-experimental Single group	18 university students from various backgrounds	IPA	<ul style="list-style-type: none"> <li>Role-playing</li> </ul>	Encourage students to revise and self-correct	Some improvements were reported on phonological awareness but no significant difference between pre-and post-tests.	N/A	N/A
21.	Najima, Kato, Tamura, and Yamamoto (2021)	To encourage students to improve speaking proficiency	Interview Single group	10 university students from Japan	<ul style="list-style-type: none"> <li>ASR</li> <li>SDS</li> </ul>	<ul style="list-style-type: none"> <li>Conversation</li> <li>Question &amp; answer</li> <li>Negotiation for meaning</li> </ul>	Enjoy repeating speaking tasks	The study demonstrated the effectiveness of the system for remote learning and concentration level.	N/A	N/A
22.	Tegos, Demetriadis, and Tsiatsos (2014)	To use MentorChat to trigger constructive interaction amongst students	Post-task Questionnaire Focus group	30 Russian university student studying English	ITS	<ul style="list-style-type: none"> <li>Conversation</li> <li>Collective discussion</li> </ul>	Engage students in conversations	The 'weak'-directed agent intervention mode seems to foster a more constructive discourse among students.	N/A	N/A
23.	Tai (2022)	To examine the influence of IPAs on English as a foreign language students'	Interview	89 college freshmen from China	IPA	<ul style="list-style-type: none"> <li>Oral proficiency tests</li> <li>Interview</li> </ul>	Students provided high-quality oral input	The findings indicate that the usage of an IPA outside of class	N/A	N/A

(continued on next page)



Table 2 (continued)

No.	Authors and Year	Focus of Study	Research Design	Participants	Technological Integration	Task Design	Students' Engagement	Learning objectives	Technological Limitation	Novelty Effect
		spoken competence outside the classroom.						significantly enhanced the oral competency of EFL students		
24.	Timpe-Laughlin, Sydorenko, and Dombi (2022)	To examine the utility of spoken dialogue systems (SDSs) for EFL learning	Survey Single group	47 EFL tertiary level students in the USA	SDS	<ul style="list-style-type: none"> <li>• Conversation</li> <li>• Question &amp; answer</li> <li>• Negotiation for meaning</li> </ul>	Verbal, gestural, and physical interaction	SDS tasks are effective for practicing speaking and collecting speaking performance data	Speaking too fast or interrupting	N/A
25.	Wu et al. (2020)	To identify emotional issues in EFL IPA user experience	Interviews Between group	32 participants from a European university	<ul style="list-style-type: none"> <li>• IPA</li> <li>• Audiolingual</li> </ul>	<ul style="list-style-type: none"> <li>• Question &amp; answer</li> <li>• Negotiation for meaning</li> <li>• Multiturn conversation</li> </ul>	Inclined to participate in well-orchestrated speaking practice Bi-directional communication	EFL speakers prefer visual feedback to support their interaction which boosts their confidence A dialogue system that is embedded in MI to improve student engagement in collaboration tasks	Difficult to recognize students' utterances	N/A
26.	Xie, Liu, Chen, and Liu (2021)	To incorporate a motivational online conversational agent (MOCA) to enhance learner engagement in computer-supported collaborative learning (CSCL)	Survey Between group	40 university students from China	RAI & GAI				Lack of seamless integration affects students' experience	N/A
27.	Yin and Satar (2020)	To explore the frequency and patterns of NfM language learning	Between group	8 undergraduate university students from China	CALL	Negotiation for meaning	Engage in negotiation for meaning	Low-level students benefit the most, and higher-level students are dissatisfied with the systems.	Not able to understand Out-of-context response	Yes
28.	Zou, Liviero, Hao, and Wei (2020)	To use an AI dialogue system to support students' speaking skills in EAP courses	A mixed-method	113 Y <sup>1</sup> and Y2 university students from China	CALL	<ul style="list-style-type: none"> <li>• Pronunciation</li> <li>• Task-based conversation</li> <li>• Presentation</li> </ul>	Students preferred AI as lack of teacher feedback	Positive attitudes toward the AI dialogue systems for speaking skills	Poor voice recognition Limited prosodic system	N/A

discussions with two independent specialists who brought their own expertise and perspectives to the interpretation of the data (Böhm, 2004, p. 270). By incorporating these external viewpoints, the authors aimed to minimize potential biases and enhance the credibility of the research findings (Lianto, 2019). The collaborative approach adopted during the axial coding stage facilitated a comprehensive and nuanced understanding of the subject matter. Through ongoing dialogue and negotiation, the authors and the independent specialists were able to collectively refine the categories and subcategories, ultimately reaching a consensus on the interpretation of the data (Harati et al., 2019). This iterative process not only contributed to the development of a robust analytical framework but also fostered a sense of shared ownership over the research findings (Vollstedt & Rezat, 2019a).

In the third stage of the research process, the authors engaged in selective coding with the aim of clustering the axial codes gathered during the previous stage. This step involved a thorough examination and synthesis of the axial codes to identify key dimensions, sub-dimensions and establish relationships between them, ultimately leading to the development of a cohesive framework that captured the essence of the findings. To maintain the rigor and validity of the research, the authors carried out selective coding individually, ensuring that each researcher's unique perspective and insights were taken into account (Vollstedt & Rezat, 2019a). Throughout this process, the authors held regular discussions with subject matter experts to review their progress, share interpretations, and address any discrepancies or uncertainties that arose. These collaborative efforts were vital in ensuring that the analysis was comprehensive, well-founded, and grounded in the data (Holton, 2007). The continuous exchange of ideas and the integration of expert feedback contributed to a more nuanced understanding of the subject matter over the research findings (Moghaddam, 2006). This iterative process continued until the authors and the subject matter experts reached a point of complete accord, signifying that a shared understanding and agreement had been achieved regarding the central themes, relationships between codes, and overall framework derived from the data (Vollstedt & Rezat, 2019a). By reaching this stage, the researchers ensured that the analysis accurately reflected the insights and patterns emerging from the collected data, ultimately enhancing the credibility and generalizability of their findings.

After three steps of open coding, axial coding and selective coding, we adopted dimensional analysis to identify and categorize dimensions related to interactional competence supported by AI dialogue systems for EFL at a university level. Dimensional analysis is a technique for characterizing complex phenomena in which meaning is socially established based on perspective and context according to grounded theory methodology and social interactionism (Morse et al., 2016). The purpose of dimensional analysis is to construct theory based on data analysis, including discovering and characterizing important dimensions, inferring correlations, and establishing linkages between dimensions (Bowers & Schatzman, 2016).

#### 2.4.2. Method of dimensional analysis

We followed Kools et al.'s (1996) influential three steps when apply dimensional analysis. First, we generated or identified dimensions and their qualities, and then sub-categorized them to expose tentative notions by expanding data. This portion of the preliminary study focuses on identification, since codes take the shape of AI dialogue systems for EFL. This procedure continues until a critical mass of dimensions and attributes has been attained. Second, we created an explanatory matrix and assigned higher weight to certain aspects while disregarding others (similar to constant comparison in Grounded Theory). Each dimension was elevated to this level during the process of determining the central perspective, and the dimension that provides the greatest explanation for the relationship between dimensions is ultimately selected as the central or key perspective from which to organize the data (Kools, McCarthy, Durham, & Robrecht, 1996). The remaining aspects are classified as salient, relevant, marginal, or irrelevant (Kools et al., 1996).

Third, we utilized the integrated core viewpoint to explain the patterns and interactions between the phenomenon's aspects by use of an explanatory matrix, thereby disclosing everything involved. This is not an exhaustive list, but it does include the majority of technologies that are commonly and often utilized in EFL educational research dealing with interactional competence. This paper identifies 6 dimensions out of 28 studies, such as: technological integrations ( $n = 28$ ), task design ( $n = 28$ ), students' engagement ( $n = 28$ ), learning objectives ( $n = 28$ ), technological limitation ( $n = 10$ ) and novelty effects ( $n = 3$ ).

Technology integration in language learning involves various sub-dimensions, such as automatic speech recognition (ASR), audiolingual, intellectual personal assistant (IPA), spoken dialogue system (SDS), computer-assisted language learning (CALL), dialogue management systems (DMS), cognitive immersive language learning environment (CILLE), data-driven development (D3), intelligent tutoring system (ITS), and retrieval-based AI & generative AI (RAI & GAI). In designing language learning tasks, aspects include conversation, negotiation for meaning, question and answer, and storytelling. Aspects of engagement of students includes bi-directional communication and productive language exercises. Learning objectives include academic achievement, increased concentration, enhanced cultural awareness, and learning through collaboration. Aspects of technological limitations include unnatural voice, incomplete phrases, failed communication, and gibberish outputs. Additionally, technology novelty can lead to a novelty effect. Fig. 1 illustrates a hierarchical diagram for enhancing EFL university students' interactional competence.

This paper aims to present an overview of the technological applications of dialogue systems for EFL in university education. The research articles retrieved from the databases were screened twice. First, an examination of the abstracts and titles of the retrieved material is carried out in order to determine whether these articles match the minimal inclusion criteria. Second, the entire text of the publications that were included in the study was analyzed and retrieved using the CQUniversity database retrieval tool. The databases yielded a total of 211 articles. After screening and removing duplicates and unsuitable content, the total number of articles was reduced to 180 articles. Then, after the eligibility screening procedure, a total of 115 articles were retrieved and assessed for eligibility. Then, any remaining publications that were not directly relevant to AI dialogue systems for EFL or EFL learning were excluded from this research. As a result, a total of 28 articles were chosen for this study.

### 3. Overview of EFL dialogue system for interactional competence

Early in the 1980s, academics were motivated by the insights learned from previous computer-assisted dialogue system development attempts to define new goals for dialogue systems (Tai, 2022). Underwood (1984) was the first person to advocate for the creation of communicative dialogue systems, a strategy that would use recent advances in AI to integrate meaningful communication practice in tutorial systems. Among other variables, this would be performed via verbal connection with others. FAMILIA, which was developed by Underwood in 1982, is a written Spanish conversational program. Its pattern-matching capabilities were somewhat inspired by the first chatbot, ELIZA (Underwood, 1984).

Over the last decade, technological advancements in dialogue systems for EFL seem to have merged into new systems that employ sophisticated NLP in tandem with more clear learning goals and better strategic evaluation techniques (Tai, 2022). Focusing on corrective feedback, dialogue systems like Sasha involve sophisticated conversation management and extremely granular evaluations of effectiveness. Multimodality, relevant and authentic communication settings, goal-oriented interactions, mixed initiative, and complex conversation management are already partially shown by these systems, which will be crucial for future enhancements to the EFL dialogue system (Petersen,

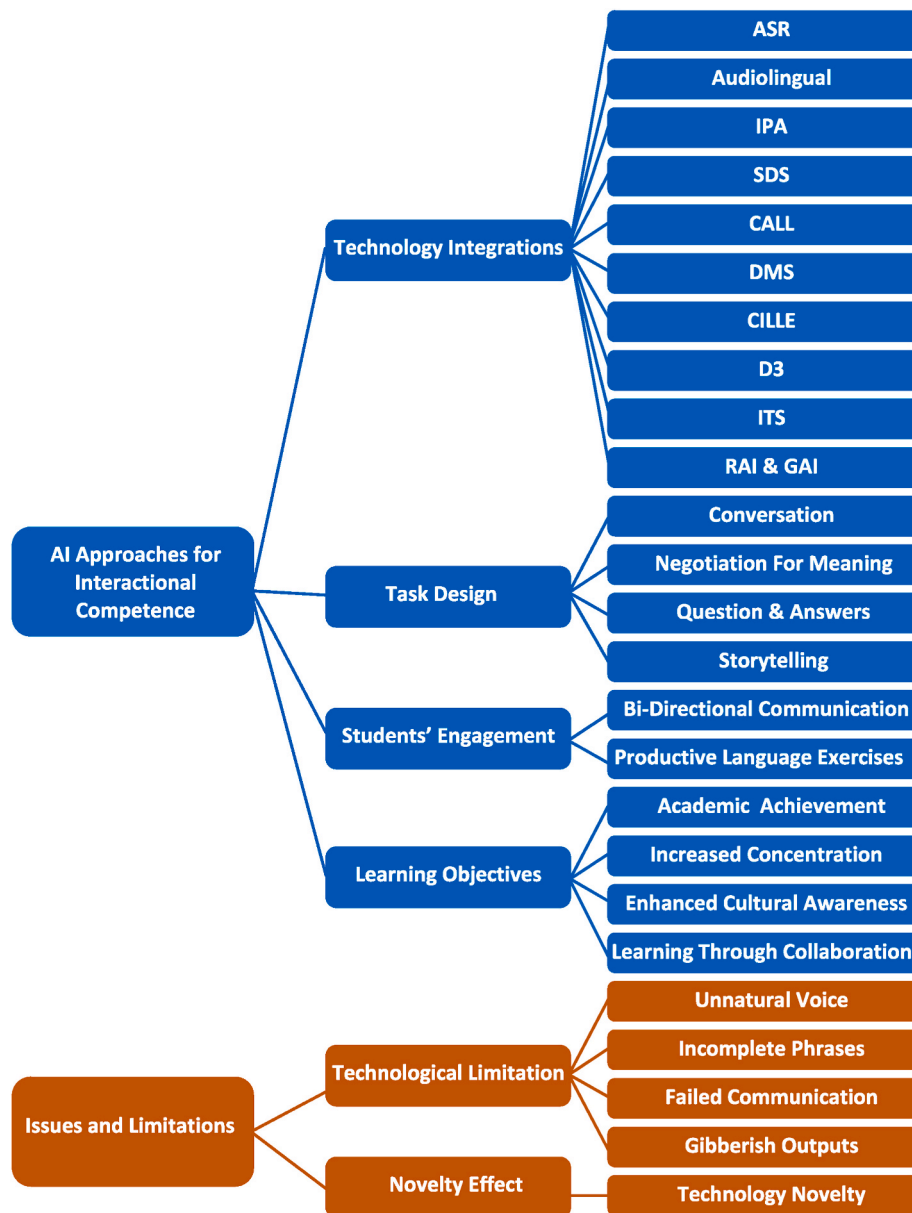


Fig. 1. A hierarchical diagram for enhancing EFL university students' interactional competence.

2010).

Recently, large-scale pre-trained models (PTMs), bidirectional encoder representations from transformers (BERT), and generative pre-trained transformers (GPT) have recently garnered enormous success and established themselves as an important milestone in the field of AI. Because of their intricate pre-training objectives and massive model parameters, large-scale PTMs are able to effectively extract information from both labelled and unlabeled data sets. PTMs are autoregressive language models that, with the use of deep learning, generate writing that is eerily similar to that produced by humans; these models have approximately 175 billion parameters.

The majority of dialogue systems for EFL is item-based, and the systems focus on the discussion as the instructional unit. This distinguishes it from production activities centered around separate items, which is similar to a sentence found in most language instructional courseware (Heift & Schulze, 2015). A conversation, as opposed to being a syntactic unit, is a pragmatic unit, containing interactional practice and numerous complex phenomena that language students are required to gain communicative competence (Heift & Schulze, 2015). Multiple

conversational turns co-construct meaning in dialogue systems. Without a set of turns, the systems cannot be considered as dialogue-based ones. Kramsch (1986) defines the term "interaction" as a communicative exchange that entails negotiating intended meanings, and modifying one's speech to the impact one aims to have on the listener. It requires defining one's own objectives, anticipating the listener's reaction and probable misunderstandings, and getting at the closest possible observation between intended, perceived, and expected meanings. Thus, interactional competence may be defined as a theory of the knowledge that participants bring to and actualize in interaction, and includes an explanation of how this information is gained (Young, 2011).

#### 4. Results

Through open coding, axial coding, and selective coding, we determined that it would be appropriate to classify and code the dimensions that were identified in the chosen studies into a six-dimension classification of EFL dialogue systems, each of which has its own objective and function in optimizing or limiting learners' interactional competence in

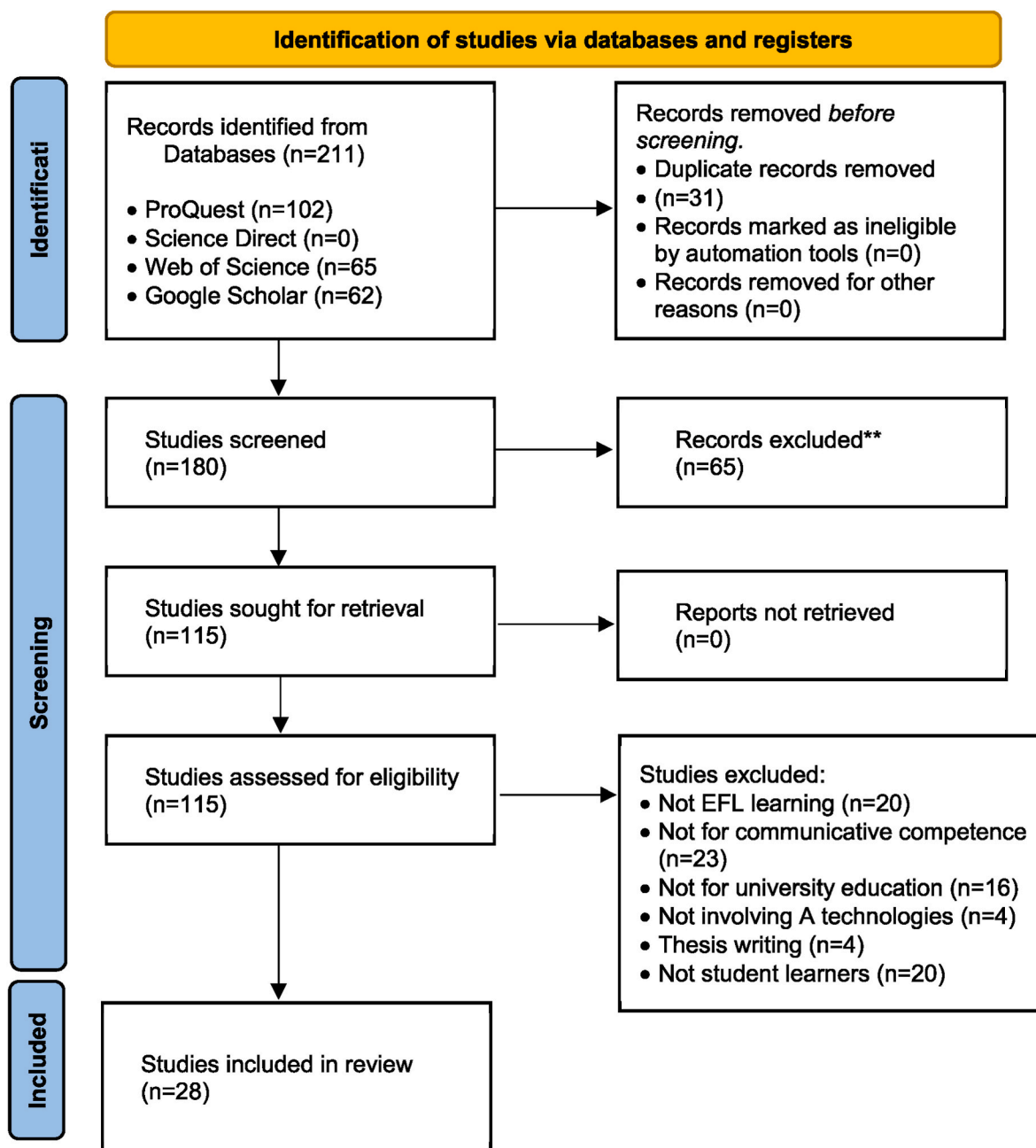


Fig. 2. The PRISMA flowchart.

university level of study. Each dimension includes multiple primary families of AI approaches. This is by no means a complete list, but it does include the majority of AI approaches that are commonly dealing with learners' interactional competence in university level of study. Table 2 presents six dimensions, including (1) technological integration, (2) task design, (3) student engagement, (4) learning objectives, (5) technological limitations, and (6) the novelty effect.

Technological integration in language learning is the use of technology to improve students' learning experience (El Shazly, 2021). Technology enhances students' opportunities for authentic connection with native speakers and other language learners of varying proficiency levels, both within and outside of the classroom (El Shazly, 2021). Task design is a systematic way in which task description, workflow and task execution plan are structured. Task design for language acquisition helps to facilitate language learning by emphasizing on the regular communication and the usefulness of the language rather than

explaining the language forms (Hismanoglu & Hismanoglu, 2011). Learner engagement is a measure of the quantity and quality of a learner's involvement in their learning (Wu et al., 2020). Learning objectives are statements that clarify what students are expected to acquire when learners have accomplished a course or programme (Morino, Lopez, & Ono, 2017, pp. 367–373). A technological limitation is the inability of computer software or hardware to perform a certain function (McTear, Callejas, & Griol, 2016). The novelty effect refers to the propensity for performance to initially improve when new technology is implemented—not because learning or accomplishment have really improved, but rather as a result of increased interest in the new technology (Fryer, Ainley, Thompson, Gibson, & Sherlock, 2017).

The finding indicates that AI dialogue systems serve a communicative purpose of EFL acquisition in university education as students find dialogue systems alleviate their speaking anxiety, enhance their vocabulary retention and offer immediate feedback to boost their



willingness to interact in the target language. However, one paper finds that AI dialogue systems increase students' anxiety instead of reducing it.

The target language in the dialogue system studies primarily focused on learning EFL. As for research design, most studies employed single-group experiments ( $n = 10$ ), followed by survey ( $n = 6$ ), between-groups ( $n = 5$ ), a mixed-method ( $n = 3$ ), and questionnaire ( $n = 3$ ) (see Table 2). With regards to task designs, most studies employed conversation strategy ( $n = 12$ ), followed by negotiation for meaning strategy ( $n = 10$ ), questions & answers strategies ( $n = 6$ ), and pre and post-experiment design ( $n = 2$ ) (see Fig. 3). A single-group experiment is a type of quasi-experiment in which the result of interest is assessed only once after the exposure of a non-random set of participants to a particular intervention. The purpose of this experiment is to measure the impact of the intervention, such as to evaluate a training program (Ryu & Cheong, 2017). A between-group study design is an experiment that employs two or more groups of participants, each of whom is put through a different testing process simultaneously (Ryu & Cheong, 2017). A pre- and post-experiment design is a method of research in which participants are asked to complete the same assessment tasks both before and after they are given intervention or are subjected to a condition. These tasks are used to determine whether or not there are any changes in the participants' performance that could be attributed to the treatment or condition (Ryu & Cheong, 2017). A survey study is a means of gathering information or data as reported by participants (Misje, Bosnes, Gåsdaal, & Heier, 2005). A questionnaire study is a research instrument that consists of a set of questions that is used to gather meaningful information from respondents. The structure of these instruments is similar to that of an interview, and it consists of questions that can be responded to either written or spoken (Misje et al., 2005). A mixed-method is a research approach whereby quantitative and qualitative data are collected within the context of the same study (Pluye, Gagnon, Griffiths, & Johnson-Lafleur, 2009).

Two major technical grounds were discovered among the dialogue system studies: technologies for establishing human–dialogue system relationships and embedded cognition via AI content design. The first theoretical grounding was the interactive relationship between humans and the AI dialogue system. An AI dialogue system is able to autonomously communicate with students from various educational backgrounds (Moussalli & Cardoso, 2021, p. 226) by fostering students' readiness toward EFL interactional practices and also enhancing their

pronunciation and vocabulary (Ayedoun, Hayashi, & Seta, 2019). Additionally, the processes of dialogue systems can alleviate students' speaking anxiety (Bashori et al., 2020). Via the audiolingual method, dialogue systems identify and adapt individual students' interactive behaviors to successfully engage students in the target language use (Hanif, 2016). This theoretical viewpoint in social psychology demonstrates a similar common basis in learning (Bashori et al., 2020). In order to apply this approach to dialogue systems, it is vital for dialogue systems to exhibit similar dimensions and knowledge. This increased language engagement, enhanced speaking abilities, and increased desire and interest in learning. Nevertheless, novelty effects have been recorded as hierarchically structured activity designs of dialogue systems are inclined to leave little variation for students to respond impromptu (Fryer et al., 2017). Table 2 contains information on the articles that were chosen, such as the names of the authors, the focus of the studies, research design, participants in the study, implantation duration, technological integration, task design, learning objectives, and technical limitations.

The second technique basis was applying automatic speech recognition (ASR) cognition through dialogue-based content design. Dialogue-based content design involves acoustic, lexicon and language models (Lestari & Nugraha, 2017). After the ASR system was initially trained, the lexicon and a corpus of audio data are leveraged as input to generate the language and acoustic models. Audio content will be displayed on the screen consisting of tangible human-like dialogue that performs natural interactions (Lestari & Nugraha, 2017).

#### 4.1. Technological integrations

Technological integration refers to the use of technology for improving students' learning experience (El Shazly, 2021). This principle is particularly important in language learning, where technology has been shown to increase student engagement and motivation, resulting in improved learning outcomes. By effectively integrating AI technologies, students can develop essential skills such as critical and creative thinking, collaboration, communication, and problem-solving, which are essential for success in the 21st century. Therefore, incorporating technology is crucial in enhancing student learning and success (Pham & Sampson, 2022; Wu et al., 2022; Xia, 2020). The technology behind chatbots includes AI, natural language processing (NLP), and machine learning (El Shazly, 2021). There are various technologies involved in

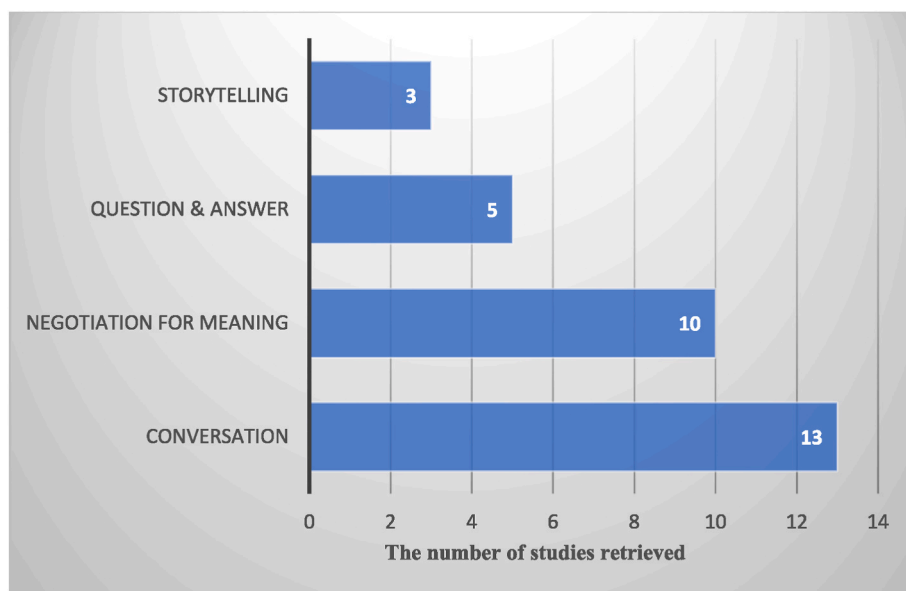


Fig. 3. Task design of dialogue systems.

improving students' pronunciation, aiming for successful spoken interaction, e.g., ASR, audiolingual, IPA, SDS, CALL, DMS, CILLE, D<sup>3</sup>, ITS, and RAI & GAI.

Several studies used more than one language training approach at the same time in the design of interactional practice activities. The most often used method was automatic speech recognition (ASR) (n = 8), followed by audiolingual (n = 5), intelligent personal assistant (IPA) (n = 5), spoken dialogue systems (SDS) (n = 5), computer-assisted language learning (CALL) (n = 3), dialogue management systems (DMS) (n = 2), cognitive immersive language learning environment (CILLE) (n = 1), data-driven development (D<sup>3</sup>) (n = 1), intelligent tutoring system (ITS) and the retrieval-based AI & generative AI (RAI & GAI) (n = 1). Table 3 illustrates the technological integration embedded in the dialogue system, which includes Automatic Speech Recognition, Audiolingual, Intellectual Personal Assistant, Spoken Dialogue System, and Computer-Assisted Language Learning.

ASR is a cutting-edge technology that enables machines to convert speech signals into text or commands after speech recognition. Audiolingual incorporates a drill routine of listening and speaking in language acquisition. Via the audiolingual method, dialogue systems identify and adapt individual students' interactive behaviors to successfully engage students in the target language use (Hanif, 2016). IPA is a software agent that performs tasks or provides services for a user's instructions or inquiries. CALL is a method of teaching and learning in which content learned is presented, reinforced, and assessed using computer and computer-based resources. SDS is a computer system capable of conversing with a user. It consists of two essential components: a voice recognizer and a text-to-speech component. DMS are the NLP components responsible for analysing and contextualising user-to-chatbot interactions. D<sup>3</sup> is a method of software development that uses data to direct the development process. ITS refers to computer learning environments that assist students in acquiring information and skills. ITS adopts powerful intelligent algorithms that adapt to the learner on a fine-grained level and execute complicated learning concepts. The retrieval-based AI uses keyword matching to choose the most relevant answer in both machine and deep learning, whereas generative based AI uses predefined responses to generate new dialogue based on the conversational training data.

Eight studies have embedded ASR in the dialogue systems. There are two main components of ASR for EFL acquisition: (1) the ability to distinguish accented from incorrect pronunciation and (2) the ability to offer the relevant evaluation on the quality of pronunciation. These studies show that the dialogue systems improved students' pronunciation and vocabulary, and students' vocabulary retainment and conversation skills significantly improved.

Five studies have adopted audiolingual function, whereby AI dialogue systems were designed to improve EFL students' interaction and oral communication. Findings show that dialogue systems were able to identify and adapt individual students' interactive behaviors to successfully engage students in the target language use. Five studies have

adopted an IPA dialogue system for the speaking and listening development of EFL English students. Findings show that students' interactional competence was considerably improved with IPA task design. Five studies included SDS in the dialogue system, and the findings showed that dialogue systems enhance students' performance in speaking. Two studies included CALL in the dialogue systems, and findings showed that low-level EFL students benefit the most, and their speaking proficiency improved dramatically after interacting with the system. CALL reinforces and evaluates the content that has to be taught. It often consists of a substantial number of interactive components.

Two studies adopted DMS, and the findings showed that dialogue systems motivate EFL students' willingness to communicate. One study introduced a CILLE in which an AI dialogue system enables multi-modal interactions to better immerse language learning. The dialogue system in CILLE provides a sense of being unintrusive when supporting multi-party interactions. One study has adopted D<sup>3</sup> in the dialogue system, Enskill®English, to foster foreign language learning, and the findings suggested that Enskill®English is useful for acquiring spoken English abilities. One study included ITS in the dialogue system to trigger constructive interaction among students, and the results found that the 'weak'-directed agent intervention mode seems to foster a more constructive discourse among students. One study, as illustrated in Table 3, embedded RAI & GAI in the dialogue system to enhance learner engagement in computer-supported collaborative learning, and findings indicated that the dialogue system improved student engagement in collaboration tasks.

Several studies revealed that language technological integration in the design of interactional practices in dialogue systems for university education place a considerable emphasis on ASR (Chen et al., 2020; Dizon, 2020; Najima et al., 2021) and audiolingual systems (Wu et al., 2020). Based on the literature, the ASR system provides students with additional opportunities for creating output in the target language. Most EFL students praised the diversity of task activities and the ease with which the ASR can be accessed. Students believed that the ASR integrated system could foster their interactional competence and enhance vocabulary memorization and retention. Moreover, the AI dialogue system embedded with ASR provides an excellent learning atmosphere for EFL students who feel self-conscious about speaking in public. Several studies showed that the ASR-based dialogues systems could provide a low-anxiety environment to foster students' speaking. One of their preferred speaking strategies is the innovative negotiation for meaning strategies in EFL acquisition. Table 4 provides brief review of findings on the interactional practice dialogue system for university education in alignment with research questions.

#### 4.2. Task design

Task design for language acquisition is a method facilitating language learning that emphasizes communication and the usefulness of the language rather than explaining the language forms. Language students are encouraged to complete tasks that require to utilize the foreign language or second language (Hismanoglu & Hismanoglu, 2011). Task design is critical for language learning because well-designed tasks can promote effective language learning outcomes (Yu & Chen, 2018). Chen, Bear, Hui, Santhi-Ponnusamy, and Meurers (2022) claim that language learning tasks should be designed in a way that promotes meaningful and authentic communication, engages learners, and provides opportunities for practice and feedback. Effective AI task design should also be aligned with learning goals and objectives and take into account learners' interests, backgrounds, and proficiency levels. In digitally-mediated learning activities, AI task design is particularly crucial in terms of goal setting, feedback, and balancing skill difficulty. By paying careful attention to task design, language teachers can ensure that language learning tasks are engaging, relevant, and aligned with learning objectives, thus promoting effective language learning outcomes (Liang, Yang, Shan, & Kim, 2021; Zhang, Zhao, Ouyang, & Zhao,

**Table 3**  
Technological integration embedded in dialogue system.

Approach	Authors
Automatic Speech Recognition	Chen et al. (2020), Divekar et al. (2021, pp. 1–29), Dizon (2017), Dizon (2020), El Shazly (2021), Mohammadzadeh and Sarkhosh (2018), Morino et al. (2017, pp. 367–373), Najima et al. (2021)
Audiolingual	Dizon (2017), El Shazly (2021), Kim (2017), Li and Peng (2021), Wu et al. (2020)
Intellectual Personal Assistant	Chen et al. (2020), Dizon and Tang (2019), Hsu et al. (2021), Moussalli and Cardoso (2021), Wu et al. (2020)
Spoken Dialogue System	Divekar et al. (2021); Mohammadzadeh and Sarkhosh (2018), Morino et al. (2017, pp. 367–373), Najima et al. (2021), Timpe-Laughlin et al. (2022)
Computer-Assisted Language Learning	Mazzilli (2021), Yin and Satar (2020), Zou et al. (2020)

**Table 4**  
Positive cognitive communication skills, and learning objectives.

Type of Cognition	Contributing Elements to EFL Students' Cognitive Progression in University Education
Retention	Interactions with the AI dialogue systems for EFL that are reinforced by multi-modal stimuli on certain vocabulary items
Understanding	EFL interaction with the dialogue systems leads to increased concentration
Social-norms	Students have the opportunity to get a deeper cultural awareness, and their intercultural competence is also enhanced
Application	Rephrasing commands when the system did not understand their utterances
Analysis	Asking questions, repeating or generating words or phrases verbally, and conversing in role-playing scenarios
Communication skills	Contributing Elements to EFL Students' Communicative Skills in University Education
Interactional practice	Shadowing practicing oral skills and mimicking a natural conversation
Pronunciation	Emphasize lexical items and sentence patterns in conversations
Vocabulary usage	Vocabulary memorization, and everyday dialogues
Story-retelling	Enhancing students' productive communication and creative skills
Negotiation for meaning	Involving asking for clarification, rephrasing, and verifying
Emotional state	Contributing Elements to EFL Students' Emotional States with AI Dialogue Systems
Enjoyment	Enjoyed interacting with AI Dialogue systems for EFL as they are non-judgmental
Relaxed	Feel less anxious in a relaxing environment
Motivated	Highly motivated in learning English with the use of an AI EFL dialogue system
Frustrated	Technological issues and unnatural TTS sound

2016).

The general functionalities from the reviewed studies include conversation ( $n = 13$ ), negotiation for meaning ( $n = 10$ ), question & answers ( $n = 5$ ), and storytelling ( $n = 3$ ). The negotiation for meaning method is believed to improve comprehension skills in the language acquisition process, during which interactivity plays a vital role in producing comprehensible input for students, asking for clarification, and providing an opportunity to restructure modified output (Dizon, 2020). A conversation function in a dialogue system is able to record, understand, and analyze vocal or text input and respond in a chosen language from a user (Xie et al., 2021). The question-and-answer method in dialogue systems is a style of content delivery, and it consists of a chatbot asking questions and the students answering those questions (Najima et al., 2021). Storytelling in dialogue systems is an innovative method that employs computer technology to enhance language learning, and the technique assists the development of language and digital abilities in students (Morino et al., 2017, pp. 367–373).

Several studies showed that conversation practise helps language students to assimilate their learned information, and integrate a range of cognitive abilities to produce meaningful spoken communication. Mohammadzadeh and Sarkhosh (2018) integrated conversation function in an AI dialogue system to improve students' speaking abilities, and the results showed that the dialogue system enhances students' performance in speaking. Wu et al. (2020) carried out a study on university EFL students' attitudes towards an IPA dialogue system in relation to interactional practice. The findings revealed that EFL speakers placed a significant emphasis on adjustments intended to increase the possibility of being intelligible. This may be due to students' speech limitations in the target language. Although students enjoyed the interactional practices with a dialogue system, students felt frustrated as they had to exert a considerable effort to rephrase commands when the system did not understand their utterances.

Ten studies entailed negotiation of meaning strategy in language learning, which is a process through which speakers arrive at a mutually understanding of what has been said. Divekar et al. (2021, pp. 1–29) presented an immersive cognitive language learning environment embedded with negotiation for meaning functions to engage students in a non-dyadic multi-modal environment which makes the EFL learning non-intrusive. The findings revealed that students' proficiency in speaking increased dramatically between pre-test and post-test. EFL students were also exposed to cultural elements, which enriched their understanding of the target language. In the immersive dialogue system, students observed and experienced unfamiliar social practices and norms, via negotiation for meaning techniques students have the opportunity to get a deeper cultural awareness, and their intercultural competence is also enhanced. Students indicated that cultural absorption fostered their willingness to communicate.

Few studies included question and answer design in the dialogue systems, Najima et al. (2021) adopted the question-and-answer method to improve students' speaking proficiency, and findings indicated that the dialogue system is effective for language learning and further revealed that the students' eye gaze behaviors changed between question-answering tasks. The eye gaze behaviors act as an indicator of the level of concentration of the students. Three studies integrated storytelling in the dialogue system to provide a sense of familiarity and trust. The storytelling method engages the listener with the story and makes them more receptive to learning. Morino et al. (2017, pp. 367–373) investigated the effectiveness of digital storytelling on Japanese university EFL students' attitudes and awareness toward CALL, and the findings showed that storytelling increased students' awareness of CALL and improved their speaking proficiency.

Studies showed that the task designs were aligned with language learning methods, such as negotiation for meaning and storytelling techniques (Timpe-Laughlin et al., 2022; C. H. Wu et al., 2022). In the literature, it shows that task design such as negotiation for meaning embedded in synchronous computer-mediated communication has been extensively studied (Wu et al., 2020), and two additional stages, confirmation and reconfirmation, have been extended to the process of negotiation routine (Timpe-Laughlin et al., 2022). At the confirmation stage, students confirm with their counterparts that the issue is attended to or unattended through a reaction to response which signals the students' acceptance or remaining unaccepted. At the reconfirmation stage, it indicates the closure of the negotiation routine signaling students' understanding (Timpe-Laughlin et al., 2022).

#### 4.3. Students' engagement

Learner engagement is a measure of the quantity and quality of a learner's involvement in their learning (Wu et al., 2020). Student engagement is essential for effective and practical language learning experiences (Tao et al., 2023). One of the effective ways to enhance student engagement is by incorporating AI through personalized and adaptive learning systems that can cater to the individual student's needs, interests, and proficiency levels. This view is supported by Diwan, Srinivasa, Suri, Agarwal, and Ram (2023) who claimed that students are more likely to participate in peer-to-peer learning, and collaboration facilitated by AI. When students are fully engaged in their studies, students tend to seek out additional learning opportunities when they are fully engaged in their studies, leading to better language learning outcomes. The higher that the students are engaged in their learning, the more motivated, interested, and committed they are towards their studies. Furthermore, student engagement promotes a positive learning environment where students feel supported, challenged, and valued through the assistance of AI. This positive AI environment, in turn, fosters a sense of belonging and enhances learning outcomes (Diwan et al., 2023; Thomas, Sarma, Gajula, & Jayagopi, 2022; Yang, Chen, Flanagan, & Ogata, 2022).

Several studies show that students are more inclined to participate in

well-orchestrated speaking practice using a contextualized and applicable way to the dialogue (Divekar et al., 2021, pp. 1–29; Wu et al., 2020; Xie et al., 2021). In addition to facilitating bi-directional communication (Xie et al., 2021) with verbal, gestural, and physical interaction to enable students to practice receptive (such as understanding words and language) (Timpe-Laughlin et al., 2022) and productive (such as actively speaking and writing) language skills (Divekar et al., 2021, pp. 1–29), language instructors were present to provide procedural and learning support to students throughout the interactive tasks. In most cases, students participated in productive language exercises such as asking questions, repeating or generating words or phrases verbally, and conversing in role-playing scenarios.

Language instructors assisted them with a variety of tasks, such as game introduction and feedback, and students were also able to learn from their experience to improve their learning outcomes (Bashori et al., 2020). Xie et al. (2021) incorporated an immersive learning environment called OpenSim to improve student engagement, and the system also tracks individual students' study curves. Bashori et al. (2020) claimed that EFL students speaking anxiety was dramatically reduced via interacting with the dialogue system, and their engagement was enhanced in the anxious-free environment. Timpe-Laughlin et al. (2022) highlight that EFL students enjoyed the interaction with the dialogue system with speaking activities as they received immediate feedback.

#### 4.4. Learning objectives

Cognitive learning objectives from interactional practice with a dialogue system were associated with significant (a) academic achievement (Morino et al., 2017, pp. 367–373), such as understanding, application and analysis, (b) increased concentration (Najima et al., 2021), such as retention, (c) enhanced cultural awareness (Divekar et al., 2021, pp. 1–29) such as social-norms and 4) learning through collaboration (Divekar et al., 2021, pp. 1–29) such as association.

Learning objectives are a critical component of effective learning because they provide a clear description of what students are expected to achieve as a result of their instruction. They guide the development of effective plans for instruction, activities, and assessments (Al Braiki, Harous, Zaki, & Alnajjar, 2020). Incorporating AI in learning objectives can help ensure that students acquire the necessary knowledge and skills to succeed in today's fast-paced technological environment. With clear learning objectives, students can better focus their efforts and achieve their learning goals, leading to successful application of their newly acquired knowledge and skills (Wong, Ma, Dillenbourg, & Huan, 2020). The communicative abilities of students, such as interactional practice, pronunciation, vocabulary usage, story-retelling, and negotiation for meaning were significantly improved (Morino et al., 2017, pp. 367–373). In particular, the speed at which students speak and their response accuracy is enhanced (Wu et al., 2020), such as grammatical (Manda, Nurlaila, & Indri, 2017), and lexical correctness (Bashori et al., 2020). Table 4 highlights the impact of dialogue systems on students' emotional state, feelings of enjoyment, relaxation, and motivation.

#### 4.5. Technological limitation

Although a dialogue system can assist students in their language acquisition, it is vital to recognize the limits of their technical competence. The apparent unnaturalness of the computer-generated voice, which students compared to humans, was the most commonly reported technical obstacle (McTear et al., 2016). AI continues to advance in language learning, it is crucial to understand its limitations to enable learners to make informed decisions about the use of specific technologies for supporting their learning goals (Bresciani, Dabić, & Bertello, 2022, p. 102063). Zhai and Wibowo (2022) argue that understanding technological limitations allows learners to avoid unrealistic expectations and frustration, and instead focus on leveraging AI technology effectively. By understanding the limitations of AI, learners can identify

the most appropriate AI tools and resources to support their language learning needs, and avoid wasting time and resources on technologies that may not be effective or appropriate for their language learning context (Ayedoun et al., 2019; Divekar et al., 2021, pp. 1–29).

When students input incomplete phrases, failed communication was also discovered, or gibberish outputs were produced from dialogue systems (Ayedoun et al., 2019). Students' positive emotional states (e.g., curiosity) in language acquisition tend to diminish with the absence of emotion and visual indications during interactions (El Shazly, 2021). Dialogue systems with weak AI were unable to interpret students' inputs. The dialogue system, for example, is unable to distinguish new subjects provided by students because their thoughts veer in unforeseen ways as the interaction progresses (Divekar et al., 2021, pp. 1–29). Moreover, they are isolating language students from the language learning environment due to dialogue systems' odd robotic voices and incapacity to have lengthy conversations. Divekar et al. (2021, pp. 1–29) observed that students expressed frustration with technological glitches and experienced exhaustion after absorbing the instructional content.

#### 4.6. The novelty effect

A novelty effect occurs when the results of a research study are influenced by the novelty or newness of a particular intervention or product being used. This effect can make it difficult to determine whether the outcomes of the study are actually effective or just a result of the novelty factor (Fryer et al., 2017). AI dialogue systems are not immune to this effect, and it has been noted that learners may experience short-term gains in engagement and performance due to the novelty effect. This novelty effect may temporally boost students' motivation or learning performance due to the technology novelty (Fryer et al., 2017). According to Fryer et al. (2017), this effect occurs with the dialogue system during a 16-week experimental study. Students' interest in speaking tasks reported decreased slowly after the first communication task. Similarly, Lubis, Sakti, Yoshino, & Nakamura (2019) argued that the dialogue system was perceived as a curiosity by the students rather than as a long-term collaborator in their everyday language practice, and delivering a workshop before the first session may help lessen the novelty impact. All studies suggested that a prolonged period of the experiment will diminish the novelty effect, and in this way, students will benefit from the advantages of technological innovation of EFL acquisition. To minimize the impact of the novelty effect in AI-based dialogue systems, teachers and researchers can consider delivering a workshop or training session before using the technology to familiarize learners with the AI and integrate it as a long-term collaborator in language practice. This way, students can benefit from the advantages of AI innovation for EFL acquisition over a prolonged period of time (Lubis et al., 2019).

### 5. Discussion

#### Q1: The key evaluation dimensions of AI dialogue systems for enhancing EFL in the university

Six dimensions were identified in the study which includes (1) technological integration, (2) task designs, (3) students' engagement, (4) learning objectives, (5) technological limitation, and (6) the novelty effect, as shown in Table 2. The finding indicates that AI dialogue systems serve a communicative purpose of EFL acquisition in university education as students find dialogue systems alleviate their speaking anxiety, enhance their vocabulary retainment and offer immediate feedback to boost their willingness to interact in the target language. However, one paper finds that AI dialogue systems increase students' anxiety instead of reducing it.



### 5.1. Technological integrations

There are various technologies involved in improving students' pronunciation, aiming for successful spoken interaction, several studies used more than one language training approach at the same time in the design of interactional practice activities, five studies have adopted audiolingual function, whereby AI dialogue systems were designed to improve EFL students' interaction and oral communication. Two studies adopted DMS, and the findings showed that dialogue systems motivate EFL students' willingness to communicate. Several studies, as summarized in Table 3, place a considerable emphasis on ASR to enhance students' interactional competence.

However, not all studies alleviate EFL students' speaking anxiety, enhance their vocabulary retainment and offer immediate feedback to boost their willingness to interact in the target language. El Shazly (2021) argues that foreign language anxiety has been a recurring issue in the domain of EFL acquisition. In his study, the author investigated the role of the AI dialogue system embedded with ASR in speaking practice from 48 Egyptian university students enrolled in an EFL course. Over eight weeks of a quasi-experimental study, he found that students suffered dramatically from foreign language anxiety, notably in their speaking capabilities which were at similar percentages between pre-test and post-test measurements. This situation may be derived from concerns about failing the course instead of fostering students' interests in EFL acquisitions. EFL students' affective status has been compromised by the high-stake study under the academic context. Thus, students' final results may have inflicted severe constraints on their self-image, preventing students from engaging in communicative activities. El Shazly (2021) also argues that students' anxiety levels were not alleviated when being exposed to a non-threatening AI dialogue system, and this finding contrasts with the literature that AI dialogue systems positively enhance students' engagement with non-judgmental AI dialogue systems (El Shazly, 2021).

On the other hand, technological integration of an audiolingual method is used to develop bottom-up language skills, such as word recognition (Ayedoun et al., 2019). In word association exercises, students were able to gain receptive language skills, such as vocabulary and phrases. In addition, students participated in constructive language usage by administering story-retelling tasks. Such engagement enhanced students' productive communication (e.g., spoken communication abilities) and creative skills (Morino et al., 2017, pp. 367–373). However, based on the studies reviewed, the audiolingual method, which is also known as rote learning, is the most vital component in acquiring new vocabulary. A few studies (Mei, 2018; Saydaliyeva, Atamirzayeva, & Dadaboyeva, 2020) found that repetition and imitation lead to limited linguistic development in university education. As the audiolingual method is mechanical associative learning, it ignores constructivist language learning, which emphasizes collaboration and idea exchanges (Saydaliyeva et al., 2020). Studies find that university EFL students are inclined to construct their own linguistic knowledge of the material related to their own experiences. Learning is enhanced when university students are encouraged to establish views and develop their thoughts about the linguistic materials being studied (El Shazly, 2021; Morino et al., 2017, pp. 367–373; Saydaliyeva et al., 2020).

### 5.2. Task design

Task design is a systematic way in which task description, workflow and task execution plan are structured. Task design for language acquisition helps to facilitate language learning by emphasizing on the regular communication and the usefulness of the language rather than explaining the language forms (Hismanoglu & Hismanoglu, 2011).

Based on the literature, language acquisition can be achieved by adopting storytelling techniques for such purposes as (a) learning vocabulary meaning related to the content of a story, (b) formulating individualized questions related to the story, and (c) repeating certain

linguistic dimensions of the actors' lines (Alsadoon, 2021). EFL students were willing to participate in interactive language learning activities such as storytelling and role-play using communicative, meaning-based language learning approaches (Divekar et al., 2021, pp. 1–29). Although it is important to note that students' spoken output tended to be closed-ended answers at both the lexical and syntax levels, this indicates that future efforts should be directed toward developing tasks that emphasize intelligibility in order to fulfil meaning-focused instruction.

In this systematic review, studies (Dizon & Tang, 2019; Morino et al., 2017, pp. 367–373) find that the integration of storytelling in AI dialogue systems is construed as fitting technologically with task-based, project-based, and thematic strategies, and this is because the storytelling technique is consistent with critical and dialogic approaches of learning, in which space is made for collective thinking, experimentation, and peer review (Morino et al., 2017, pp. 367–373). In Morino et al.'s (2017, pp. 367–373) study, authors expect students to craft a narrative in a bilingual form because it prevents the compartmentalization of languages and it acknowledges students' interrelationships between what is in mind and what is in usage. This makes it possible for students' bilingual or multilingual repertoires to be represented in a holistic manner and for syncretic identities to be built. This helps stimulate university EFL students' confidence and encourages them to place a higher value on their bilingual or multilingual skills. Yin and Satar (2020) found that the lack of non-linguistic elements, such as paralinguistic gestures and facial expressions, imposes an undue burden on written disclosure, delineating more creative clarity. This type of forced output in synchronous computer-mediated communication, according to Swain's (1995) Comprehensible Output Hypothesis, forced output increases students' awareness of the linguistic characteristics of their interlanguage. When students encounter a gap in their linguistic knowledge of the target language, learning occurs (Yin & Satar, 2020).

Based on the retrieved studies, task design, such as everyday conversation and storytelling, are still prevalently used in EFL learning. Yet, EFL students in university education are required to gain a body of information from various subject domains and disciplines. Thus, problem-solving skills are needed for students to cultivate self-directed learning skills. Problem-solving skills encompass a wide range of constructivist perspectives and stretch over many fields as university students are able to attend to issues from a real-world perspective. Students are able to construct new linguistic knowledge with a more profound comprehension and better retention outcomes (Azman & Shin, 2012).

Thus, a well-designed EFL dialogue system for university education can have more linguistic content through various techniques such as debate, discussion or problem-solving. For example, Google's Dialogflow endorses users to modify conversational content by alternating pre-set datasets (e.g., intents) (Sagar et al., 2021). BotStar, an online dialogue system platform, enables users to create conversational flows by dragging and dropping intents on a designed dashboard (Shevat, 2017). These functions allow EFL instructors to script students' learning experiences in accordance with the desired learning goals (Sagar et al., 2021; Shevat, 2017).

### 5.3. Students' engagement

Learner engagement is a measure of the quantity and quality of a learner's involvement in their learning (Wu et al., 2020). Multiple studies demonstrate that students are more likely to engage in well-organized speaking practice that is contextualized and related to the dialogue system (Wu et al., 2020). In most cases, students participated in productive language exercises such as asking questions, repeating or generating words or phrases verbally, and conversing in role-playing scenarios.

However, no studies have introduced or mentioned with what kinds of technology were embedded to foster EFL students' engagements so as to enhance to learning outcomes. One suggestion might be the integration of humor, empathy practice and culture awareness in the dialogue

system to encourage students in a culturally humorous yet empathetic environment. Palmer and Menard-Warwick (2012) argue that humor, empathy practice and culture awareness in class plays a significant role in increasing the levels of psychological and emotional wellness which in return help students achieve higher levels of academic accomplishment. The English language has been identified as a means of intercultural communication between individuals of varied linguistic and cultural origins (Zhang, Mandl, & Wang, 2011).

Humor, the ability to recognize or communicate what is humorous, is both a source of amusement and a technique of dealing with tough or embarrassing circumstances and stressful occurrences (Thomas, Nelson, & Silverman, 2015). To have a sense of humor requires not just a comprehension of the language and words, but also of how they are used, what they signify, the underlying culture, the connotations, and the unwritten messages (Thomas et al., 2015).

Empathy is described as the capacity to identify and understand the feelings and thoughts of other people, as well as the ability to project one's own thoughts and feelings onto those of another (Matsuhisa et al., 2021). Empathy is the language that allows language learners to speak with a level of depth that fosters connection and enables them to comprehend the logic behind another person's decisions, regardless of how divergent they may be to their own (Matsuhisa et al., 2021).

Previous research has shown that the use of humor in the classroom helps enhance students' interactional skills (Gonulal, 2021; Hismanoglu, Ersan, & Turan, 2018), and improves students' understanding and retention of learning material (Farnia & Mohammadi, 2021; Gonulal, 2018). Pomerantz and Bell (2011) advocated for the inclusion of humor not simply as a method for boosting enjoyment, but also as an essential component of the process of acquiring practical interactional abilities. The authors argue that when it comes to verbal information, such as words, sentences, and visual information, learner's memories are improved through the use of humor. As a result of factors such as greater attention and enhanced memory rehearsal, the humor effect can be ascribed to improved encoding and retrieval of funny information as contrasted to—and often at the price of—non-humorous information.

The perception of humor engages both hemispheres of the brain. The left hemisphere is the “logical brain,” which is responsible for language competence, logical analysis, and detailed humor recognition, while the right hemisphere is the “creative brain,” which is responsible for imagining and comprehending humor (Palmer & Menard-Warwick, 2012). As humor is universal yet tinted by culture (Palmer & Menard-Warwick, 2012), in an egalitarian framework, students from horizontal collectivist cultures emphasize sociability and interdependence with others and practice affiliative humor. Meanwhile, students from vertical collectivist cultures comply to the hierarchical social order stratified by their in-groups and are willing to conduct altruistic actions for their in-groups, they prefer to use self-defeating humor. This is in contrast to a low individualist society (collectivism), in which values such as self-sacrifice, interdependence, and generosity towards others are emphasized (Palmer & Menard-Warwick, 2012). In addition, Members in collectivist cultures tend to focus more on others than on themselves, resulting in a society with a far deeper sense of empathy and care for others (Palmer & Menard-Warwick, 2012).

Regarding empathy, it is the capacity to perceive the experiences of others from various cultures, including both high context and low context societies (Matsuhisa et al., 2021). A person forms and defines his or her sense of self by conforming to certain society norms, values, and beliefs. Students from high-context cultures are expected to discover the hidden complexity from their culturally ingrained intuitions, whereby they perceive low-context communication as overly detailed. Individuals who appreciate a low-context communication style, on the other hand, see high-context conversationalists as lacking in transparency or incapable of effectively transmitting ideas (Matsuhisa et al., 2021).

#### 5.4. Learning objectives

Cognitive learning objectives from interactional practice with a dialogue system were associated with significant academic achievement (Morino et al., 2017, pp. 367–373), and to achieve better scores for certain English tests seems to be an utterly important goal for most university EFL students. English tests such as the International English Language Testing System (IELTS), Cambridge English Test, and Pearson's Test of English (PTE) Academic, have adopted communicative approaches with supervised interactions between the interviewer and the interviewee or interactions between different interviewees (Mastura, 2021). Thus, interactional competence becomes an important component of the curriculum in EFL education because they allow students to interact effectively with others in the target language (Timpe-Laughlin et al., 2020). By developing interactional competence in handling linguistic queries, language learners gained greater leverage to optimize their academic achievement (Mastura, 2021).

Interactional competence help to increase concentration and collaborations (Najima et al., 2021), this is similar to Young (2011)'s statement that the author claimed that interactional competence is distinct from communicative competence in that it places a greater emphasis on interactions on the development of spoken English. Conversational competence is very context-dependent and co-constructed by speakers who participate in the conversation. As a result, it requires a larger focus and concentration on the relationships that occur between interlocutors. The capacity to deploy interactional resources (turn-taking, repair, boundaries, speech actions, etc.) using accessible language resources as required by the speaker or the hearer in order to convey their communicative intents in real situational circumstances is what we mean when we talk about interactional competence (Galaczi & Taylor, 2018).

Interactional competence also enhances cultural awareness (Divekar et al., 2021, pp. 1–29), which is in par with Hymes' (1972) statement, the author argued that not only does competence refer to an individual's knowledge of the forms and structures of language, but it also extends to how an individual uses language in actual social and cultural situations. Hymes outlined four categories of information that people draw upon while interacting in various social and cultural settings. These categories include what is conceivable to accomplish with language, what is practicable to do with language, what is proper to do with language, and what is really done according to the social norm.

#### 5.5. Technological limitation

Although a dialogue system can assist students in their language acquisition, it is vital to recognize the limits of their technical competence. The apparent unnaturalness of the computer-generated voice, which students compared to humans, was the most commonly reported technical obstacle (McTear et al., 2016).

There are three obstacles in using the system. First, the inability of an AI dialogue system to interact with students of heavily accented speech, and thus students' interactional competence may be compromised (Chen et al., 2020). Interactional practice is arguably the most critical component of EFL development (Najima et al., 2021). Many language students may be unable to engage in AI dialogue interaction due to the fact that some AI dialogue systems appear to have difficulty fully understanding EFL speech with non-standard English variants. This highlights a broader problem with AI-assisted dialogue, and speech recognition accuracy across various language variants, take English dialects as an example, Tatman (2017) finds that YouTube automatic captioning systems struggle to perform its accuracy with Scottish English. Johnson (2019) states that the choices of the dialogue system are too limited and restricted, and encountered bugs, and the author believed that natural language understanding was designed to understand the simple language of students. Thus, suggestions proposed by Divekar et al. (2021, pp. 1–29) were that language students should be

cognizant of technical restrictions, and this is because dialogue systems operate by using pre-determined subjects and pre-stored phrases, which affect natural language processing. Thus, students should try to avoid colloquialisms and partial sentences.

Second, unnaturalness of the computer-generated voice was reported technical obstacle (McTear et al., 2016), which tends to refrain EFL learners participation of interaction with the dialogue system. Usually, synthetic speech is often created from text, and due to a revolution in AI, the audio quality of synthetic speech has significantly improved over the last several years. Elmers, Werner, Muhlack, Möbius, and Trouvain (2021) suggest that a breath noise is embedded into the system to reduce the unnaturalness of the computer-generated voice which showed that there is an improvement in students' recollection for sentences that were preceded by a breath noise as opposed to sentences that did not include a breath noise. The authors also found that a significant influence for phrase length, arguing that shorter sentences had more accuracy for recollection than longer ones when breath noises were included.

Third, dialogue systems with weak AI were unable to interpret students' incomplete inputs, when students' ideas wander unpredictably as the engagement proceeds (Divekar et al., 2021, pp. 1–29). The inability to interpret students inputs also isolate language learners from the language-learning environment as a result of dialogue systems cannot extend further discussions. Divekar et al. (2021, pp. 1–29) report that students suffered weariness after digesting instructional information and showed irritation with the dialogue system cannot interpret students' inputs. Suggestions to increase the response accuracy from incomplete input can follow most popular classification approach to use imputation to replace incomplete values with plausible values (Elmers et al., 2021). Tran, Zhang, Andreae, Xue, & Bui (2018) believe that in many real-world datasets, it is impossible to prevent incomplete inputs or missing values, and authors provided novel methods for merging imputation, clustering, and feature selection for classification with partial data in order to enhance efficiency without sacrificing precision. Clustering is used to decrease the number of examples utilized by the imputation. Feature selection is used to eliminate duplicate and unnecessary characteristics of training data, which significantly decreases the cost of imputation. The research also studies Differential Evolution's (DE) capacity to find feature subsets with inadequate data. The results demonstrate that integrating imputation, clustering, and feature selection not only improves classification accuracy, but also significantly decreases the computation time necessary to estimate missing values when classifying new instances.

Another recommendation to increase the response accuracy from incomplete inputs is to remove stop words and linking synonyms by using Natural Language Toolkit (NLTK), and the phrase "stop word" refers to frequent phrases that do not add to any deeper meaning and that the dialogue systems are trained to disregard (Kaur & Buttar, 2018). Wordnet's definition of synonyms is that they are "words that convey the same notion and are interchangeable in various situations." As a result of this, synonyms are categorized into sets that are not arranged in any particular order (synsets). These synsets are what we make use of in order to generate the synonyms and antonyms that are shown in the following programs. When stop words are removed, the low-level information is eliminated, which allows the dialogue systems focus on key features of the output. Eliminating stop words does not only reduce the vector space, but also improves performance overall.

### 5.6. The novelty effect

The novelty effect is a phenomenon that can occur in studies involving new or innovative technologies, such as dialogue systems, where the initial excitement and curiosity generated by the technology can lead to enhanced engagement and performance among learners (Hammad & Bahja, 2023; Miguel-Alonso, Rodriguez-Garcia, Checa, & Bustillo, 2023). However, it is important to note that the gains made by learners may be short-term, and the novelty effect can mask the true

effectiveness of the technology (Fryer et al., 2017). Therefore, researchers and educators must be cautious in interpreting the results of studies involving novel technologies and consider the potential impact of the novelty effect on the outcomes. The study by Lubis et al. (2019) highlights the potential negative impact of the novelty effect on the use of AI dialogue systems for language practice. Learners may perceive the technology as a curiosity rather than a long-term collaborator, limiting their ability to fully engage with the technology and benefit from its features (Fryer et al., 2017). Therefore, to minimize the impact of the novelty effect on the outcomes of studies involving new technologies, appropriate training and support should be provided to learners. Engagement with AI dialogue systems can be sustained by making it meaningful and useful for students, and overcoming the novelty effect that can cause engagement to fall once the initial excitement wears off. Tsay, Kofinas, Trivedi, and Yang (2020) suggests that gamified learning systems with gamified elements can attract users initially, but engagement can fall once the novelty effect wears off. Longitudinal iterative cycles are important for observing the novelty effect and generating ways to overcome its potential negative impact on engagement. The study found that for a substantial proportion of students, engagement with the gamified learning system was intrinsically driven, which helped to overcome the novelty effect. This resulted in engagement becoming habitual, playful, and game-like, and helped to sustain engagement over time.

### 5.7. Summary of the six dimensions

Most studies indicated that technological integration in dialogue systems for EFL improves students' interactional competence for both within and outside of the classroom. As task design in the dialogue system systematically delivers task description, and organises workflow and task execution plan by emphasizing on frequent communication and the utility of the language rather than teaching the language forms to aid language learning. Then, using learner engagement as a tool to quantify the extent and quality of a learner's participation in their EFL, as well as to elucidate the knowledge and skills that students have attained from a certain course or programme. As the development and implementation of an AI dialogue system in EFL is still in its infancy stage, technological limitations occur. These limitations include the inability of computer software or hardware to perform a certain function, as well as the novelty effect, which is when learners' performance initially improves when new technology is implemented as a result of increased interest in the new technology.

This study has examined 28 articles on the interactive design of dialogue systems for EFL for university education. They include technological integration, task design, students' engagement, and learning objectives. Based on the study, there is overwhelming support for English language learning in university education. Students acquire declarative and procedural knowledge through constant repetition or rehearsal of linguistic activities, and they get personalized support in a mistake-friendly atmosphere via a dialogue system, which enhances students' language abilities without time and space constraints (Divekar et al., 2021, pp. 1–29). When students get an opportunity to apply their linguistic skills verbally in a simulated real-life environment, it reduces their anxiety and enhances their desire to converse in the target language (Timpe-Laughlin et al., 2020). Nonetheless, this type of language learning engagement seems to be diminishing significantly over time as students' interest fade away due to a novelty effect, the lower value of dialogue systems compared to human facilitators (Bashori et al., 2020). However, challenges still remain in relation to offering students tangible assistance to become expressive and eloquent in speaking discourse (Divekar et al., 2021, pp. 1–29).

From the literature, AI dialogue systems in EFL learning in university education are used to offer synchronous formative feedback to students about their spoken performance and to stimulate their metacognitive reasoning strategies by analyzing students' learning patterns (Dizon,

2020). The studies found that the majority of students are from countries where English is a foreign language, and countries such as China, Japan, and South Korea embrace a vertical collectivist culture that conforms to the hierarchical social system stratified by their in-groups and are prepared to perform altruistic acts for their in-group. This contrast with students from horizontal collectivist cultures emphasizes sociability and interdependence with others in an egalitarian context (Zhang et al., 2011).

Regardless of whether learning takes place within and outside of the classroom, EFL students can get access to essential information through dialogue systems. Forsythe (2017) referenced what Young (2011) stated in the situated learning theory that learning takes place in various places and times. Students have access to an abundance of language support materials and authentic situations or activities via mobile phones and the Internet. On the other hand, An, Gan, and Wang (2020) argue that students may regulate their own pace through dynamic and spontaneous self-regulated learning based on cognitive, emotional, motivational, and behavioral components. In An et al.'s (2020) self-regulated learning theory, the authors believed that self-generated ideas, emotions, and behaviors are geared towards the achievement of a learner's own objectives. Such elements enable students to adapt their behaviors to meet their learning goals in a variety of educational contexts. Consequently, students may manage and adapt their knowledge and plans to be suitable with the ever-changing contexts or settings for learning. In foreign language acquisition, An et al. (2020) state that language students may achieve more success in their language learning and usage if they are more strategic in their efforts.

Q2: Gaps arising from the results, additional necessary study, and recommendations to achieve effective systems

To increase EFL students' long-term learning outcomes, instead of short-term knowledge acquisition for just proficiency test, the integration of humor, empathy practice and culture awareness in class plays a significant role in increasing the levels of psychological and emotional wellness which in return help students achieve higher levels of academic accomplishment and interactional competence (Palmer & Menard-Warwick, 2012). It has been discovered that 28 studies seemed overlooked the cultural backgrounds, emotional recognition, and humor aspects in the design of dialogue system for EFL aimed for enhancing university students' interactional competence. The study also found that the importance of debate and problem-solving skills seemed was not considered in the design of AI dialogue system, thus additional necessary studies are needed in order to achieve effective dialogue systems for EFL.

### 5.8. Lack of cultural considerations

Despite the fact that the learner's cultural background plays a role in the formation of the EFL learner's identity as well as the cognitive characteristics a learner adopts, none of 28 studies have considered the importance of learners' cultural backgrounds in designing the dialogue system for EFL. According to Benfilali, Nadif, Khartite, Benattabou, and Bouih (2021), acquiring a second language is a way for a person to exhibit their culture as well as their cultural identity. A method to EFL learning incorporates intercultural language learning, which encompasses helping students acquire an awareness of their own culture as well as an identification with other cultures. In terms of cognitive characteristics, Mehrotra and Yilmaz (2015) contend that the cognitive characteristics coexist in the development of language proficiency. The first cognitive characteristic is an analytical procedural system that is rule-based, and the second cognitive characteristic is a declarative system that is example-based. Adults are more likely to be defined by the former, whilst children are defined by the latter when learning English as a foreign language. Young people naturally have a greater potential for learning than their adult counterparts do, therefore they have an advantage when it comes to the process of natural language acquisition.

This is because of the numerous cognitive qualities that young people have. These cognitive traits have an effect on how an EFL student takes in information, how that knowledge is perceived, and how it is organized in their minds.

The use of culturally relevant humor in EFL class is proven to improve the long-term outcomes and wellbeing. Chabeli (2008) argues that cross-cultural humor enhanced the emotional intelligence and critical thinking of learners. In an egalitarian framework, students from horizontal collectivist cultures emphasize sociability and interdependence with others and practice affiliative humor. Whereas students from vertical collectivist cultures comply to the hierarchical social order stratified by their in-groups and are willing to conduct altruistic actions for their in-groups, they prefer to use self-defeating humor. This is in contrast to a low individualist society (collectivism), in which values such as self-sacrifice, interdependence, and generosity towards others are emphasized (Palmer & Menard-Warwick, 2012). Critical thinking is a sort of cognitive capacity that is essential for relatively rapid and practical learning, and it entails the development of flexibility, which aids in the analysis and evaluation of knowledge after it has been acquired. Learners who have acquired good critical thinking skills are often in a continual state of inquiry, striving to establish logical explanations of their experience and to negotiate their prior knowledge (Thomas, 2011). It is advised that incorporating culturally relevant comedy models is beneficial for establishing a dynamic learning environment that enables students to make fair judgements and build critical thinking abilities.

To achieve an effective dialogue system for EFL, this paper recommends that it is vital to include cultural considerations in the behavioral models of dialogue systems. The content and the form of dialogue systems' utterances aimed toward certain cultures must be consistent. For instance, depending on the culture simulated by the dialogue system, various dimensions for a small chat could be chosen, and different discourse markers can be deployed to express politeness. Based on the diverse cultural backgrounds, the volume and quality of motions may also vary. The perception and selection of communication behaviors are heavily influenced by often overlooked cultural factors (Endrass & André, 2014).

If cultural dimensions are not taken into consideration while developing interactive behaviors for a dialogue system, this may cause an adverse effect on user acceptability. Several studies, such as Shibani et al. (2015), showed on how virtual agents were perceived racial distinctions in their verbal and nonverbal communication. Researchers altered the behavior of the virtual agent to mimic the behavior of people of various races, but the agent's appearance remained racially undefined throughout the project. Researchers analyzed students' views of the virtual agent's ethnicity by having them complete a questionnaire. Students were able to appropriately classify the virtual agent to various ethnicities based on its behavioral manners. They tried to reflect their ethnicity in a way that can be used for educational purposes. Even if culture is not explicitly considered during the modelling process of a virtual agent, the character will nevertheless exhibit characteristics of a particular cultural background. This is usually that of the programmer since they are the one who determines whether or not the agent's behavior is natural. When a programmer comes from a Western culture where direct eye contact is thought to reflect honesty, this may seem to be unacceptable to a person from a different cultural background, such as a member of Asian culture, and this behavior may be considered unfriendly by others.

### 5.9. Lack of empathetic considerations

Most studies found that dialogue systems were able to identify and adapt individual students' interactive behaviors to engage students in the target language use. However in EL Shazly's (2021) study, the author investigated the role of the AI dialogue system embedded with ASR in speaking practice from 48 Egyptian university students enrolled



in an EFL course. Over eight weeks of a quasi-experimental study, he found that students suffered dramatically from foreign language anxiety, notably in their speaking capabilities which were at similar percentages between pre-test and post-test measurements. Under such circumstance, no solution, maybe empathetic encouragement, was provided in EL Shazly's study to alleviate anxiety or stress when EFL learners experience a significant negative impact during their speaking training session while using a dialogue system. According to Gennaro et al.'s (2020) findings, dialogue systems that include empathic responses have a palliative impact on individuals who are nervous. There have been various studies using empathic dialogue systems to elicit appropriate emotional responses to the emotional condition of students learning English in settings other than universities (Fryer, Coniam, Carpenter, & Lăpușneanu, 2020; Santos, Ethel, & Resurreccion, 2020). The authors believed that it is essential to include culturally related empathic modelling into dialogue systems for EFL so that cognitive and affective components can capture and react to university EFL students' emotional state regardless of where they are learning (Marx & Pray, 2011).

Thus, incorporating empathy modelling in dialogue systems for EFL would be helpful for resolving issues caused by learners from two distinct cultural backgrounds and would aid EFL learners in adjusting to new environments. According to Cundiff, Nadler, and Swan (2009), empathy encompasses cultural values, ideas, and perceptions, similar to the culturally related humor. Benattabou (2020) establishes the notion of the unity of theory and reality after examining the link between the concept of empathy and pragmatics. He then subdivided empathy into pragma-linguistic value and socio-pragmatic value within the field of intercultural communication. The author came to the conclusion that it is necessary to combat ethnocentrism, negative stereotypes, and prejudice, as well as acknowledge the equality of different cultural traditions. Therefore, including culturally appropriate empathic modelling into conversation systems for EFL will provide students a better understanding of cultures other than their own and erase psychological barriers that exist between different languages and cultural norms.

To achieve an effective dialogue system for EFL to further enhance learners' learning outcomes, and the design of the dialogue systems for EFL needs to include functions or algorithms to facilitate learners to comprehend the experiences of individuals from other cultures, which includes both high context and low context civilizations (Matsuhisa et al., 2021). By adhering to specific societal standards, values, and beliefs, an individual creates and defines his or her sense of self. Students reared in high-context cultures are expected to discern the concealed complexity from their culturally established intuitions, and they observe low-context communication as being excessively descriptive. In contrast, individuals who value a low-context style of communication characterize high-context conversationalists as lacking in transparency or incapable of properly conveying ideas (Matsuhisa et al., 2021).

### 5.10. Lack of humor considerations

As technological integration in this study focuses on primarily improving students' pronunciation; task design focuses on encouraging students to complete tasks, such as negotiation for meaning, question & answers, and storytelling; students' engagement focuses on learners' involvement in their learning, no studies have focuses on dialogue for EFL that improve students' attitude toward the dialogue systems by developing culturally relevant humor, boosting their motivation and enhancing their performance. In addition, there has been no research conducted to characterize the nature of the humor mechanisms that are used by computer interfaces. This would involve, first, reflecting the cultural background of the target language, and then, providing contextual recognition strategies that guide conversations when they are required. It is essential to include culturally relevant comedy since research on language instruction in the classroom suggests that doing so might aid boost students' emotional intelligence. As a result, it is essential to incorporate culturally relevant humor (Chabeli, 2008).

Palmer and Menard-Warwick (2012) argue that humor is universal yet tinted by culture. The capacity of the brain as a whole may be increased through the use of humor. Similarly McGhee (1983) argues that the perception of humor stimulates both hemispheres of the brain at the same time. The left hemisphere is known as the "logical brain," and it is responsible for language proficiency, logical analysis, and detailed humor recognition. On the other hand, the right hemisphere is known as the "creative brain," and it is responsible for imagining and understanding humor. Learners who participated in activities that required convergent thinking showed a tendency to engage both hemispheres of the brain while working on a creation task, and learners who participated in activities that required humor exhibited brain activity that was comparable to that of learners who engaged in convergent thinking activities (Thomas et al., 2015).

To achieve an effective dialogue system for EFL, it is recommended that future research devote considerable effort to finding the impact of cross-cultural humor on the language ability of EFL students. Thomas et al. (2015) find that increased in learners' performance and good results were observed when humor approaches were utilized in the classroom. The research found that one of the most advantages of incorporating cross-cultural humor into the classroom is the improvement of learning and knowledge recall abilities. Baker, D'Mello, Rodrigo, and Graesser (2010) claimed that the continuity of learning, which includes cognitive and emotional states, happens in a particular context. The authors suggested that the cognitive and emotive states of a learner are crucial when culture-related humor is used effectively during learning. Learners are likely to boost their self-motivation and be actively engaged in their studies if the learning material and goal are entertaining and engaging (Hoa, 2011).

Based on the study carried out by El Shazly (2021), students suffered dramatically from foreign language anxiety, notably in their speaking capabilities which were at similar percentages between pre-test and post-test measurements. To encourage and motivate weak EFL students, we recommend that intercultural empathy functions are needed as cultural empathy intends to improve a relationship or to understand communication effectively, empathy contains shared emotional responses (affective) and perspective-taking (cognitive) activities. When the context is insinuated, it enables one to step into the shoes of others based on the constituents of their internal frame of reference or worldview, such as beliefs, motivations, fears, and aspirations (Krzmaric, 2010).

The use of AI algorithms in cultural elements raises several ethical and legal challenges. It requires lawmakers and media platforms to take steps to regulate the transmission of cultural and creative content to prevent discrimination or isolation. Data used as inputs to the algorithms include racial, gender, and stereotypes, which can have serious repercussions. Ethics frameworks have recently been proposed to encourage accountable, ethical, and accommodating algorithms of all stakeholders. Metcalf and Crawford (2016) argue that ethical frameworks must establish effective regulations as various social platforms so far cannot self-regulate.

### 5.11. The importance of debate and problem-solving

Through discussion or debate, the EFL students acquire greater understanding through shared viewpoints and ideas (Mei, 2018). However, some technological integrations, such as audiolingual method cannot provide the effectiveness of discussion as it is a mechanical associative learning, as most studies in this systematic review primarily focused on interactional practices such as basic vocabulary memorization and everyday dialogues. It is believed that debate is a potentially useful educational technique for EFL speaking tasks since it can scaffold the learning process in university education. During a debate stage, EFL students take up the majority of attention and are required to impromptu elaborate their ideas and viewpoints over the 'clash' period of the debate (Mei, 2018). Debates improve students' interactional

competence, in addition, their language, vocabulary, delivery, and diction will improve so as to present their arguments concisely. Debates help students to listen attentively and explain their ideas clearly (Mei, 2018).

## 6. Conclusion

This paper presented an overview of the technological applications of dialogue systems for EFL in university education and explored EFL students' interactional practices among various agents for EFL acquisition. A systematic review of existing studies on AI dialogue systems was conducted systematically between 2013 and 2022. In particular, the study questions focus on technological integrations, task designs, students' engagement, and learning objectives. The results of this systematic review indicate that the instructional design of dialogue systems makes use of communicative language learning, such as audiolingual approaches and narrative, as the most prominent language learning techniques, which led to a positive skill-based cognitive learning outcome.

### 6.1. Theoretical and practical implications

This study has both theoretical and practical contributions. From theoretical perspective, this systematic review can serve as a foundation for learning new languages and other technological domains. Following concerns from stakeholders, the current teaching of spoken English and interactional competence is still viewed as unsatisfactory. This is due to the fact that many graduates still lack the ability to communicate in a foreign language, which prevents them from participating fully in transcultural and global society (Beshir & Yigzaw, 2022; Darmajanti, 2017). The study has provided a profound understanding of the obstacles and advantages arising from the interactional practices process with dialogue systems in EFL interactional competence. Through the use of the dialogue systems, this review has presented an opportunity to perform more research that would aid in the improvement of the learning experience of students in general. Recommendation to improve response accuracy from incomplete inputs can follow the most common categorization strategy and utilize imputation to replace missing values with reasonable values (Tran et al., 2018), and using NLTK to eliminate stop words and connect synonyms can further enhance the accuracy of responses to partial inputs. When stop words are omitted, low-level information is discarded, allowing conversation systems to concentrate on crucial output characteristics. Not only does eliminating stop words lower the vector space, but it also increases overall performance.

Practically, this research sheds light on the obstacles of interactional practice in employing a dialogue system in EFL acquisition. Prior to designing a dialogue system to maximize students' outcomes, it is critical to consider spoken discourse methods, instructional design, the roles of dialogue systems, and skill-based cognitive learning techniques. The study's findings can be utilized by educators to build effective strategies for dialogue systems. The results are also useful for language instructors to better identify the situation associated with using new technology in order to provide suitable ways for enhancing students' learning experiences and outcomes.

A few relevant teaching suggestions for on-site teaching teachers are also presented below:

- Incorporate the AI dialogue system into their daily teaching routine. This system can be used to assist students in finding information, providing instant feedback and guidance, and answering routine questions on the class discussion forum. This approach can free up teachers' time, allowing them to concentrate on fostering students' critical thinking and problem-solving skills, while students can analyze problems and access necessary information efficiently. By using the AI dialogue system consistently over time, teachers can

identify students who need additional support and provide them with personalized guidance.

- Use the AI dialogue system to record and track various learning characteristics of each learner comprehensively. This will enable teachers to gain insight into each student's prior knowledge, specific learning style, preferences, and needs, and create personalized teaching and learning strategies. This approach will result in a more effective and engaging learning experience for each student.
- Incorporate the AI dialogue system to explore and pursue their passions, leading to increased engagement and motivation. Teachers can utilize the AI dialogue system to incorporate students' interests into the lessons, providing relevant hands-on activities, and offering related learning resources with multimedia-rich content such as videos, simulations, and games. This combination will help students better understand complex subjects and retain information more efficiently.

### 6.2. Limitations and future research possibilities

Existing information and analyses are limited to interactional practice for EF acquisition, which is one of the limitations of this research. In addition, the investigation is limited to material published between January 2012 and August 2022. Future research should investigate the use of alternative technologies to enhance the interactional practices of other businesses, such as the medical and hospitality sectors.

There is also a paucity of research regarding the cultural backgrounds of students while providing personalized feedback. In general, the significance of variety in feedback while considering users' cultural background is neglected. Feedback contexts and cultures are interconnected, and research on student feedback cultures is mostly influenced by cultural variance and shows students' difficulties with learning, critical thinking, and academic writing. Thirdly, research demonstrates that students suffer dissatisfaction and stress throughout the learning process (Ayedoun et al., 2019; Divekar et al., 2021, pp. 1–29), but no sympathetic approach is offered to ease these issues. In overall, the preceding research demonstrate that dialogue systems have a major favorable effect on the speaking instruction of students.

According to the findings of the study, several studies have been conducted on the technical usage of dialogue systems for EF acquisition. Future study should enhance the learning experiences of students with respect to the other three macro abilities, namely listening, reading, and writing. This study's focus is limited to an investigation of conversation systems in EFL acquisition. Age and gender should be the focus of future study, and ethnography will assist in the creation of successful learning strategies.

None of the papers included in this study was conducted longitudinally. There was a wide range of intervention periods in the experiments reviewed, ranging from a few minutes to a few weeks, which might have resulted in a novelty effect in the language learning process. Studies spanning two or more academic years should be investigated in order to determine whether students' interest or motivation in engaging with dialogue systems increases or remains the same.

Finally, the majority of the existing studies on dialogue systems for language learning relied on self-reported questionnaires for their data collection, which did not adequately illustrate the potential outcomes of utilizing dialogue systems on students' language success.

### 6.3. Recommendations

This paper has examined the use of dialogue systems in interactional practice for EFL acquisition for university education. It is found that interactional practice for university education has not yet reached the stage of mainstream use of dialogue systems in language learning. The following suggestions are made for designing an EFL dialogue system to improve student's experience:

- A prolonged period of an experiment will diminish the novelty effect, and in this way, students will benefit from the advantages of technological innovation in EFL acquisition.
- Further studies are required to develop speaking activities for EFL students which are accessible and user-friendly at educational institutions in university education for academic purposes.
- To guarantee that speaking activities are more effective, it is necessary to build a framework for deploying dialogue systems for university education.

### Declaration of competing interest

The authors have no relevant financial or non-financial interests to disclose.

### References

- Adenle, Y. A., Chan, E. H., Sun, Y., & Chau, C. (2020). Exploring the coverage of environmental-dimension indicators in existing campus sustainability appraisal tools. *Environmental and Sustainability Indicators*, 8, Article 100057. <https://doi.org/10.1016/j.indic.2020.100057>
- Al Braiki, B., Harous, S., Zaki, N., & Alnajjar, F. (2020). Artificial intelligence in education and assessment methods. *Bulletin of Electrical Engineering and Informatics*, 9 (5), 1998–2007. <https://doi.org/10.11591/eei.v9i5.1984>
- Al-Oqaily, E. I., Salam, A. R. H., & Na, K. S. (2022). The use of blackboard in the practice of English-Speaking skills among Saudi EFL learners during COVID-19. *Arab World English Journal*. <https://doi.org/10.24093/awej/covid2.23>
- Alsadoon, R. (2021). Chatting with AI bot: Vocabulary learning assistant for Saudi EFL learners. *English Language Teaching*, 14(6), 135–157. <https://doi.org/10.5539/elt.v14n6p135>
- An, Z., Gan, Z., & Wang, C. (2020). Profiling Chinese EFL students' technology-based self-regulated English learning strategies. *PLoS One*, 15(10), Article e0240094. <https://doi.org/10.1371/journal.pone.0240094>
- Arung, F., Rafli, Z., & Dewanti, R. (2019). In E. Setiawati, et al. (Eds.), *A new substantive theory for the mastery of foreign language speaking skills: A grounded research* (p. 231).
- Ayedoun, E., Hayashi, Y., & Seta, K. (2019). Adding communicative and affective strategies to an embodied conversational agent to enhance second language learners' willingness to communicate. *International Journal of Artificial Intelligence in Education*, 29(1), 29–57. <https://doi.org/10.1007/s40593-018-0171-6>
- Ayedoun, E., Hayashi, Y., & Seta, K. (2020). Toward personalized scaffolding and fading of motivational support in L2 learner–dialogue agent interactions: An exploratory study. *IEEE Transactions on learning technologies*, 13(3), 604–616. <https://doi.org/10.1109/TLT.2020.2989776>
- Azman, N., & Shin, L. K. (2012). Problem-based learning in English for a second language classroom: Students' perspectives. *International Journal of Learning*, 18(6). <https://doi.org/10.18848/1447-9494/CGP/v18i06/47648>
- Baker, R. S., D'Mello, S. K., Rodrigo, M. M. T., & Graesser, A. C. (2010). Better to be frustrated than bored: The incidence, persistence, and impact of learners' cognitive-affective states during interactions with three different computer-based learning environments. *International Journal of Human-Computer Studies*, 68(4), 223–241.
- Bashori, M., van Hout, R., Strik, H., & Cucchiari, C. (2020). Web-based language learning and speaking anxiety. *Computer Assisted Language Learning*, 1–32. <https://doi.org/10.1080/09588221.2020.1770293>
- Benattabou, D. (2020). Helping EFL students avoid socio-pragmatic failure: Focus on nonverbal intercultural competence. *TESOL and Technology Studies*, 1(1), 23–41.
- Bener, E. (2021). Flipping EFL classes in higher education: A systematic review. *Language Education and Technology*, 1(2), 90–109, 0000-0002-8587-8538.
- Benfilali, I., Nadif, B., Khartite, B., Benattabou, D., & Bouih, A. (2021). Cross gender oral communication from biological difference and socialized identity to mutual understanding. *Journal of World Englishes and Educational Practices*, 3(5), 13–27. <https://doi.org/10.32996/jweep.2021.3.5.2>
- Beshir, M., & Yigzaw, A. (2022). Students' self-repair in EFL classroom interactions: Implications for classroom dynamics. *Asian-Pacific Journal of Second and Foreign Language Education*, 7(1), 1–15. <https://doi.org/10.1186/s40862-022-00153-6>
- Bibauw, S., François, T., & Desmet, P. (2019). Discussing with a computer to practice a foreign language: Research synthesis and conceptual framework of dialogue-based CALL. *Computer Assisted Language Learning*, 32(8), 827–877. <https://doi.org/10.1080/09588221.2018.1535508>
- Böhm, A. (2004). *5.13 theoretical coding: Text analysis in grounded theory*. A Companion to.
- Bowers, B., & Schatzman, L. (2016). Dimensional analysis. In *Developing grounded theory* (pp. 86–126). Routledge.
- Bresciani, S., Dabić, M., & Bertello, A. (2022). *Collaborative technological development for addressing grand challenges: Opportunities, limitations, and new frameworks*. Elsevier.
- Butler, Y. G. (2019). Assessment of young English learners in instructional settings. [https://doi.org/10.1007/978-3-030-02899-2\\_24](https://doi.org/10.1007/978-3-030-02899-2_24). *Second handbook of English language teaching*.
- Camiciottoli, B. C., & Campoy-Cubillo, M. C. (2018). In *Introduction: The nexus of multimodality, multimodal literacy, and English language teaching in research and practice in higher education settings* (Vol. 77, pp. 1–9). Elsevier.
- Canale, M., & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics*, 1(1), 1–47.
- Chabelli, M. (2008). Humor: A pedagogical tool to promote learning. *Curationis*, 31(3), 51–59.
- Chen, X., Bear, E., Hui, B., Santhi-Ponnusamy, H., & Meurers, D. (2022). Education theories and AI affordances: Design and implementation of an intelligent computer assisted language learning system. In *Artificial intelligence in education. Posters and late breaking results, workshops and tutorials, industry and innovation tracks, practitioners' and doctoral consortium: 23rd international conference, AIED 2022* (pp. 582–585). [https://doi.org/10.1007/978-3-031-11647-6\\_120](https://doi.org/10.1007/978-3-031-11647-6_120). Durham, UK, July 27–31, 2022, Proceedings, Part II.
- 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>
- Chomsky, N. (1965). *Aspects of the theory of syntax*, 1965. Cambridge, MA: MIT Press.
- Colquhoun, H. L., Levac, D., O'Brien, K. K., Straus, S., Tricco, A. C., Perrier, L., et al. (2014). Scoping reviews: Time for clarity in definition, methods, and reporting. *Journal of Clinical Epidemiology*, 67(12), 1291–1294.
- Cundiff, N. L., Nadler, J. T., & Swan, A. (2009). The influence of cultural empathy and gender on perceptions of diversity programs. *Journal of Leadership & Organizational Studies*, 16(1), 97–110.
- Da Silva, J., Fernandes, V., Limont, M., & Rauen, W. B. (2020). Sustainable development assessment from a capitals perspective: Analytical structure and indicator selection criteria. *Journal of Environmental Management*, 260, Article 110147. <https://doi.org/10.1016/j.jenvman.2020.110147>
- Darmajanti, P. (2017). The development of interactional competence of EFL learners. In *International conference on English language teaching (ICONELT 2017)* (pp. 180–183). <https://doi.org/10.2991/iconelt-17.2018.40>
- Diederich, S., Brendel, A. B., Morana, S., & Kolbe, L. (2022). On the design of and interaction with conversational agents: An organizing and assessing review of human-computer interaction research. *Journal of the Association for Information Systems*, 23(1), 96–138. <https://doi.org/10.17705/1jais.00724>
- Divekar, R. R., Drozdal, J., Chabot, S., Zhou, Y., Su, H., Chen, Y., et al. (2021). *Foreign language acquisition via artificial intelligence and extended reality: Design and evaluation*. Computer Assisted Language Learning. <https://doi.org/10.1080/09588221.2021.1879162>
- Diwan, C., Srinivasa, S., Suri, G., Agarwal, S., & Ram, P. (2023). AI-based learning content generation and learning pathway augmentation to increase learner engagement. *Computers & Education: Artificial Intelligence*, 4, Article 100110. <https://doi.org/10.1016/j.caeai.2022.100110>
- Dizon, G. (2017). Using intelligent personal assistants for second language learning: A case study of alexa. *TESOL Journal*, 8(4), 811–830. <https://doi.org/10.1002/tesj.353>
- Dizon, G. (2020). Evaluating intelligent personal assistants for L2 listening and speaking development. *Language, Learning and Technology*, 24(1), 16–26.
- Dizon, G., & Tang, D. (2019). A pilot study of Alexa for autonomous second language learning. *CALL and Complexity*, 107.
- El Shazly, R. (2021). Effects of artificial intelligence on English speaking anxiety and speaking performance: A case study. *Expert Systems*, 38(3), Article e12667.
- Elmers, M., Werner, R., Muhlack, B., Möbius, B., & Trouvain, J. (2021). Take a breath: Respiratory sounds improve recollection in synthetic speech. *Interspeech*, 3196–3200. <https://doi.org/10.21437/Interspeech.2021-1496>
- Elo, S., & Kyngäs, H. (2008). The qualitative content analysis process. *Journal of Advanced Nursing*, 62(1), 107–115. <https://doi.org/10.1111/j.1365-2648.2007.04569.x>
- Endrass, B., & André, E. (2014). Integration of cultural factors into the behavioral models of virtual characters. *Natural Language Generation in Interactive Systems*. <https://doi.org/10.1017/CBO9780511844492.010>
- Farnia, M., & Mohammadi, S. (2021). Exploring EFL teachers' and learners' perception of L2 humor: A case study of Iranian English language institutes. *Eurasian Journal of Applied Linguistics*, 7(1), 151–168. <https://doi.org/10.7592/EJHR2018.6.1.gonulal>
- Firmstone, J. (2019). Editorial journalism and newspapers' editorial opinions. In *Oxford research encyclopedia of communication*. <https://doi.org/10.1093/acrefore/9780190228613.013.803>
- Forsythe, E. M. (2017). *A qualitative case study of Japanese university students and personal smartphone use in English as a Foreign Language classes*. Northcentral University.
- Fryer, L. K., Ainley, M., Thompson, A., Gibson, A., & Sherlock, Z. (2017). Stimulating and sustaining interest in a language course: An experimental comparison of Chatbot and Human task partners. *Computers in Human Behavior*, 75, 461–468.
- Fryer, L., Coniam, D., Carpenter, R., & Lăpușneanu, D. (2020). Bots for language learning now: Current and future directions. *Language, Learning and Technology*, 24(2), 8–22.
- Galaczi, E., & Taylor, L. (2018). Interactional competence: Conceptualisations, operationalisations, and outstanding questions. *Language Assessment Quarterly*, 15(3), 219–236. <https://doi.org/10.1080/15434303.2018.1453816>
- Gennaro, M., Krumhuber, E. G., & Lucas, G. (2020). Effectiveness of an empathic chatbot in combating adverse effects of social exclusion on mood. *Frontiers in Psychology*, 10, 3061.
- Girgin, U., & Brandt, A. (2020). Creating space for learning through 'Mm hm' in a L2 classroom: Implications for L2 classroom interactional competence. *Classroom Discourse*, 11(1), 61–79. <https://doi.org/10.1080/19463014.2019.1603115>
- Gonulal, T. (2018). Investigating the potential of humour in EFL classrooms. *The European Journal of Humour Research*, 6(1), 141–161. <https://doi.org/10.7592/EJHR2018.6.1.gonulal>
- Gonulal, T. (2021). Investigating EFL learners' humorous interactions with an intelligent personal assistant. *Interactive Learning Environments*, 1–14. <https://doi.org/10.1080/10494820.2021.1974489>



- Gorjian, B. (2017). The evaluation of using computer assisted language learning (CALL) facilities in developing EFL among Islamic Azad University practitioners: The case of computer literacy. *Iranian Journal of Applied Language Studies* (Vol. 9), 35–50. Proceedings of the first international conference on language focus.
- Gunnarsdóttir, I., Davidsdóttir, B., Worrell, E., & Sigurgeirsdóttir, S. (2020). Review of indicators for sustainable energy development. *Renewable and Sustainable Energy Reviews*, 133, Article 110294. <https://doi.org/10.1016/j.rser.2020.110294>
- Guzey, S. S., Yildiz, E. A., Demir, M. C., & Aksu-Dunya, B. (2022). Impact of Edpuzzle use on the assessment and measurement course achievement. *HAYEF: Journal of Education*, 19(1), 52–61.
- Hall, J. K., Hellebrand, J., & Doehler, S. P. (2011). L2 interactional competence and development. *Multilingual Matters*. <https://doi.org/10.21832/9781847694072>
- Hammad, R., & Bahja, M. (2023). Opportunities and challenges in educational chatbots. *Trends, applications, and challenges of chatbot technology*, 119–136. <https://doi.org/10.4018/978-1-6684-6234-8.ch005>
- Hanif, H. (2016). Using Rosetta Stone software as media in teaching English vocabulary (an experimental study at SDN No. 02 Lhoksukon). *Getsempena English Education Journal*, 2(1), Article 217667. <https://doi.org/10.46244/geej.v2i1.681>
- Harati, H., Nooshinfard, F., Isfandyari-Moghaddam, A., Babalhavaeji, F., & Hariri, N. (2019). Factors affecting the unplanned use behavior of academic libraries users: Towards an axial coding pattern. *Aslib Journal of Information Management*, 71(2), 138–154. <https://doi.org/10.1108/AJIM-04-2018-0092>
- Heift, T., & Schulze, M. (2015). *Research timeline*. <https://doi.org/10.1017/S0261444815000245>
- Hismanoglu, M., Ersan, Y., & Turan, Y. Z. (2018). Turkish EFL learners' perceptions on teachers' using humor in the EFL classroom. *Online Submission*, 6(2), 284–294. <https://doi.org/10.18298/ijlet.2820>
- Hismanoglu, M., & Hismanoglu, S. (2011). Task-based language teaching: What every EFL teacher should do. *Procedia-Social and Behavioral Sciences*, 15, 46–52. <https://doi.org/10.1177/003368820303400105>
- Hoa, N. T. M. (2011). *Developing EFL learners' intercultural communicative competence: A gap to be filled* (Vol. 86). From defining EIL competence to designing EIL learning.
- Holton, J. A. (2007). The coding process and its challenges. *The Sage handbook of grounded theory*, 3, 265–289.
- Hsu, 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*, 390–410.
- Hsu, Chen, & Yu, C.-S. (2021). Proposing a task-oriented chatbot system for EFL learners speaking practice. *Interactive Learning Environments*, 1–12. <https://doi.org/10.1080/10494820.2021.1960864>
- Huang, W., Hew, K. F., & Fryer, L. K. (2022). Chatbots for language learning—are they really useful? A systematic review of chatbot-supported language learning. *Journal of Computer Assisted Learning*, 38(1), 237–257. <https://doi.org/10.1111/jcal.12610>
- Huo, J. (2022). The role of learners' psychological well-being and academic engagement on their grit. *Frontiers in Psychology*, 504. <https://doi.org/10.3389/fpsyg.2022.484325>
- Huth, T., & Betz, E. (2019). Testing interactional competence in second language classrooms: Goals, formats and caveats. In *Teaching and testing L2 interactional competence* (pp. 322–356). Routledge. <https://doi.org/10.4324/9781315177021-13>
- Hymes, D. (1972). On communicative competence, 269–293. B. Pride, & J. Holmes (Eds.). *Sociolinguistics*, 269–293.
- Inkster, B., Sarda, S., & Subramanian, V. (2018). An empathy-driven, conversational artificial intelligence agent (Wysa) for digital mental well-being: Real-world data evaluation mixed-methods study. *JMIR mHealth and uHealth*, 6(11), Article e12106. <https://doi.org/10.2196/12106>
- Insights, T. B. (2021). *English Language Learning market size by end user (corporate learners, individual users, educational institutions, & government institutions), methodology (online learning, offline/classroom learning, & blended learning), global industry analysis, share, growth, trends, and forecast 2022 to 2030*. <https://www.thebrainyinsights.com/report/english-language-learning-market-12711>
- Jafari, N., & Ansari, D. N. (2012). The effect of collaboration on Iranian EFL learners' writing accuracy. *International Education Studies*, 5(2), 125–131.
- Johnson, W. L. (2019). Data-driven development and evaluation of Enskill English. *International Journal of Artificial Intelligence in Education*, 29(3), 425–457. [0.1007/s40593-019-00182-2](https://doi.org/10.1007/s40593-019-00182-2)
- Kampittayakul, T. (2019). The role of translanguaging in improving Thai learners' interactional competence in dyadic English as a foreign language tutorial sessions. *PASAA: A Journal of Language Teaching and Learning in Thailand*, 56, 80–111.
- Kaur, J., & Buttar, P. K. (2018). A systematic review on stopword removal algorithms. *International Journal on Future Revolution in Computer Science & Communication Engineering*, 4(4), 207–210.
- Kawinkoonlasate, P. (2020). Online Language learning for Thai EFL learners: An analysis of effective alternative learning methods in response to the covid-19 outbreak. *English Language Teaching*, 13(12), 15–26. <https://doi.org/10.5539/elt.v13n12p15>
- Kim, N.-Y. (2016). Effects of voice chat on EFL learners' speaking ability according to proficiency levels. *Multimedia-Assisted Language Learning*, 19(4), 63–88. <https://doi.org/10.15702/mall.2016.19.4.63>
- Kim, N.-Y. (2017). Effects of types of voice-based chat on EFL students' negotiation of meaning according to proficiency levels. *English teaching*, 72(1), 159–181.
- Kim, H.-S., Cha, Y., & Kim, N. Y. (2021). Effects of AI chatbots on EFL students' communication skills. *Korean Journal of English Language and Linguistics*, 21, 712–734. <https://doi.org/10.15738/kjell.21.202108.712>
- Kools, S., McCarthy, M., Durham, R., & Robrecht, L. (1996). Dimensional analysis: Broadening the conception of grounded theory. *Qualitative Health Research*, 6(3), 312–330.
- Kramsch, C. (1986). From language proficiency to interactional competence. *The Modern Language Journal*, 70(4), 366–372.
- Krzmaric, R. (2010). Empathy and climate change: Proposals for a revolution of human relationships. *Future ethics: Climate Change and Apocalyptic Imagination*, 153–172.
- Kuhail, M. A., Alturki, N., Alramlawi, S., & Alhejori, K. (2023). Interacting with educational chatbots: A systematic review. *Education and Information Technologies*, 28(1), 973–1018. <https://doi.org/10.1016/j.caeai.2021.100033>
- Lestari, D. P., & Nugraha, R. R. (2017). A spoken-based question answering system for train route service using the frame-based approach. *The Sixth International Conference on Electrical Engineering and Informatics*, 1–6. <https://doi.org/10.1109/ICEEI.2017.8312384>
- Liang, F., Yang, Y., Shan, Z., & Kim, B.-S. (2021). Application of artificial intelligence technology in leisure sports course. In *2021 2nd international conference on computers, information processing and advanced education* (pp. 772–775). <https://doi.org/10.1145/3456887.3457065>
- Lianto, F. (2019). Grounded theory methodology in architectural research. *Journal of Physics: Conference Series*, 1179(1). <https://doi.org/10.1088/1742-6596/1179/1/012102>
- Li, K.-C., Chang, M., & Wu, K.-H. (2020). Developing a task-based dialogue system for English language learning. *Education Sciences*, 10(11), 306. <https://doi.org/10.3390/educsci10110306>
- Li, B., & Peng, M. (2021). *The evaluation of a blended teaching mode based on an AI language learning platform* (pp. 1017–1021). The Second International Conference on Information Science and Education. <https://doi.org/10.1109/ICISE-IE53922.2021.00320>
- Lubis, N., Sakti, S., Yoshino, K., & Nakamura, S. (2019). Positive emotion elicitation in chat-based dialogue systems. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 27(4), 866–877. <https://doi.org/10.1109/TASLP.2019.2900910>
- Manda, R., Nurlaila, H., & Indri, W. (2017). Development of web-based computer-assisted language learning in English intensive course. *IOP Conference Series: Materials Science and Engineering*, 180(1), Article 012291. <https://doi.org/10.1088/1757-899X/180/1/012291>
- Marx, S., & Pray, L. (2011). Living and learning in Mexico: Developing empathy for English language learners through study abroad. *Race, Ethnicity and Education*, 14(4), 507–535.
- Mastura, T. (2021). Assessing learners' communicative competence in different testing system. *Online Journal of Sustainability and Leadership Research*, 1(5), 154–158. <https://doi.org/10.1371/journal.pone.0233613>
- Matsuhisa, T., Takahashi, N., Takahashi, K., Yoshikawa, Y., Aomatsu, M., Sato, J., et al. (2021). Effect of physician attire on patient perceptions of empathy in Japan: A quasi-randomized controlled trial in primary care. *BMC Family Practice*, 22(1), 1–9.
- Mazzilli, F. (2021). Using Elbot to enhance conflict-solving skills in learners of German as a foreign language. *Educazione Linguistica. Language Education*, 10. <https://doi.org/10.30687/ELLE/2280-6792/2021/01/005>
- McGhee, P. E. (1983). The role of arousal and hemispheric lateralization in humor. In *Handbook of humor research* (pp. 13–37). Springer.
- McTear, M., Callejas, Z., & Griol, D. (2016). Conversational interfaces: Past and present. In *The conversational interface* (pp. 51–72). Springer. [https://doi.org/10.1007/978-3-319-32967-3\\_4](https://doi.org/10.1007/978-3-319-32967-3_4)
- Mehrotra, R., & Yilmaz, E. (2015). Terms, topics & tasks: Enhanced user modelling for better personalization. *Proceedings of the 2015 International Conference on the Theory of Information Retrieval*, 131–140. <https://doi.org/10.1145/2808194.2809467>
- Mei, Y. (2018). Comments on the audiolingual method. *International Journal of Arts and Commerce*, 7(4), 90. [https://doi.org/10.1007/978-3-319-32967-3\\_4](https://doi.org/10.1007/978-3-319-32967-3_4)
- Metcalfe, J., & Crawford, K. (2016). Where are human subjects in big data research? The emerging ethics divide. *Big Data & Society*, 3(1), 2053951716650211 <https://doi.org/10.1177/2053951716650211>
- Miguel-Alonso, I., Rodriguez-Garcia, B., Checa, D., & Bustillo, A. (2023). Countering the novelty effect: A tutorial for immersive virtual reality learning environments. *Applied Sciences*, 13(1), 593. <https://doi.org/10.3390/app13010593>
- Misje, A. H., Bosnes, V., Gåsdaal, O., & Heier, H. E. (2005). Motivation, recruitment and retention of voluntary non-remunerated blood donors: A survey-based questionnaire study. *Vox Sanguinis*, 89(4), 236–244.
- Moghaddam, A. (2006). Coding issues in grounded theory. *Issues in Educational Research*, 16(1), 52–66.
- Mohammadzadeh, A., & Sarkhosh, M. (2018). The effects of self-regulatory learning through computer-assisted intelligent tutoring system on the improvement of EFL learners' speaking ability. *International Journal of Instruction*, 11(2), 167–184.
- Morino, A., Lopez, E., & Ono, Y. (2017). *Effects of a digital storytelling project on Japanese EFL Learners' CALL attitudes and awareness of CALL tasks*. Society for Information Technology & Teacher Education International Conference. <https://www.learntechlib.org/primary/p/177309/>
- Morse, J. M., Bowers, B. J., Charmaz, K., Corbin, J., Clarke, A. E., & Stern, P. N. (2016). *Developing grounded theory: The second generation* (Vol. 3). Routledge. <https://doi.org/10.1111/j.1548-1387.2010.01143.x>
- Moscardini, A., Strachan, R., & Vlasova, T. (2022). The role of universities in modern society. *Studies in Higher Education*, 47(4), 812–830. <https://doi.org/10.1080/03075079.2020.1807493>
- Moussalli, S., & Cardoso, W. (2021). *Intelligent personal assistants and L2 pronunciation development: Focus on English past- CALL and Professionalisation*.
- Najima, T., Kato, T., Tamura, A., & Yamamoto, S. (2021). Remote learning of speaking in syntactic forms with robot-avatar-assisted language learning system. *The 24th International Conference on Text, Speech, and Dialogue*, 558–566. [https://doi.org/10.1007/978-3-030-83527-9\\_48](https://doi.org/10.1007/978-3-030-83527-9_48)
- Nirwana, N., & Suhono, S. (2022). Phonological interference in English pronunciation produced by students at senior high school (A case study of Buginese and Javanese



- students). *Anglophile Journal*, 3(1), 1–13. <https://doi.org/10.51278/anglophile.v3i1.475>
- Palmer, D. K., & Menard-Warwick, J. (2012). Short-term study abroad for Texas preservice teachers: On the road from empathy to critical awareness. *Multicultural Education*, 19(3), 17–26.
- Petersen, K. A. (2010). *Implicit corrective feedback in computer-guided interaction: Does mode matter?* Georgetown University, 10822/553155.
- Pham, S. T., & Sampson, P. M. (2022). The development of artificial intelligence in education: A review in context. *Journal of Computer Assisted Learning*, 38(5), 1408–1421. <https://doi.org/10.1111/jcal.12687>
- Pluye, P., Gagnon, M.-P., Griffiths, F., & Johnson-Lafleur, J. (2009). A scoring system for appraising mixed methods research, and concomitantly appraising qualitative, quantitative and mixed methods primary studies in mixed studies reviews. *International Journal of Nursing Studies*, 46(4), 529–546.
- Pomerantz, A., & Bell, N. D. (2011). Humor as safe house in the foreign language classroom. *The Modern Language Journal*, 95, 148–161. <https://doi.org/10.1111/j.1540-4781.2011.01274.x>
- Roever, C., & Kasper, G. (2018). Speaking in turns and sequences: Interactional competence as a target construct in testing speaking. *Language Testing*, 35(3), 331–355.
- Rosé, C. P., & Ferschke, O. (2016). Technology support for discussion based learning: From computer supported collaborative learning to the future of massive open online courses. *International Journal of Artificial Intelligence in Education*, 26(2), 660–678. <https://doi.org/10.1007/s40593-016-0107-y>
- Ryu, E., & Cheong, J. (2017). Comparing indirect effects in different groups in single-group and multi-group structural equation models. *Frontiers in Psychology*, 8, 747, 2017-25858-001.
- Sagar, R. H., Ashraf, T., Sharma, A., Goud, K. S. R., Sahana, S., & Sagar, A. K. (2021). *Revolution of AI-enabled health care chat-bot system for patient assistance* (Vol. 778). Springer. [https://doi.org/10.1007/978-981-16-3067-5\\_18](https://doi.org/10.1007/978-981-16-3067-5_18)
- Santos, O., Ethel, & Resurreccion, R. (2020). Therapist vibe: children's expressions of their emotions through storytelling with a chatbot. *Proceedings of the Interaction Design and Children Conference*, 483–494. <https://doi.org/10.1145/3392063.3394405>
- Saydaliyeva, M., Atamirzayeva, E., & Dadaboyeva, F. (2020). Modern methods of teaching English in Namangan state university. *International Journal on Integrated Education*, 3(1), 8–9. <https://doi.org/10.31149/ijie.v3i1.256>
- Shevat, A. (2017). *Designing bots: Creating conversational experiences* (Vol. 1). O'Reilly Media, Inc.
- Shiban, Y., Schelhorn, I., Jobst, V., Hörnlein, A., Puppe, F., Pauli, P., et al. (2015). The appearance effect: Influences of virtual agent features on performance and motivation. *Computers in Human Behavior*, 49, 5–11.
- Skogmyr Marian, K., & Balaman, U. (2018). Second language interactional competence and its development: An overview of conversation analytic research on interactional change over time. *Language and Linguistics Compass*, 12(8), Article e12285. <https://doi.org/10.1111/lnc3.12285>
- Siddaway, AP, Wood, AM, & Hedges, LV (2019). How to do a systematic review: A best practice guide for conducting and reporting narrative reviews, meta-analyses, and meta-syntheses. *Annual Review of Psychology*, 70, 747–770. <https://doi.org/10.1146/annurev-psych-010418-102803>
- Skogmyr, K., & Balaman, U. (2018). Second language interactional competence and its development: An overview of conversation analytic research on interactional change over time. *Language and Linguistics Compass*, 12(8), Article e12285. <https://doi.org/10.1111/lnc3.12285>
- Swain's, M. (1995). Three functions of output in second language learning. *Principles and Practice in Applied Linguistics: Studies in Honor of HG Widdowson*, 125–144.
- 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>
- Tao, X., Shannon-Honson, A., Delaney, P., Dann, C., Xie, H., Li, Y., et al. (2023). Towards an understanding of the engagement and emotional behaviour of MOOC students using sentiment and semantic features. *Computers & Education: Artificial Intelligence*, Article 100116. <https://doi.org/10.1016/j.caeai.2022.100116>
- Tatman, R. (2017). Gender and dialect bias in YouTube's automatic captions. *Proceedings of the First ACL Workshop on Ethics in Natural Language Processing*, 53–59. <https://doi.org/10.18653/v1/W17-1606>
- Tegos, S., Demetriadis, S., & Tsiatsos, T. (2014). A configurable conversational agent to trigger students' productive dialogue: A pilot study in the CALL domain. *International Journal of Artificial Intelligence in Education*, 24(1), 62–91. <https://doi.org/10.1007/s40593-013-0007-3S>
- Thomas. (2011). Developing first year students' critical thinking skills. *Asian Social Science*, 26–35.
- Thomas, B., & Skinner, H. (2012). Dissertation to journal article: A systematic approach. *Education Research International*, 2012, 1–11 (I) <https://doi.org/10.1155/2012/862135>.
- Thomas, Nelson, J. K., & Silverman, S. J. (2015). *Research methods in physical activity* (7 ed.). USA: Human kinetics Publisher.
- Thomas, C., Sarma, K. P., Gajula, S. S., & Jayagopi, D. B. (2022). Automatic prediction of presentation style and student engagement from videos. *Computers & Education: Artificial Intelligence*, 3, Article 100079. <https://doi.org/10.1016/j.caeai.2022.100079>
- Timpe-Laughlin, V., & Dombi, J. (2020). Exploring L2 learners' request behavior in a multi-turn conversation with a fully automated agent. *Intercultural Pragmatics*, 17(2), 221–257. <https://doi.org/10.1515/ip-2020-0010>
- Timpe-Laughlin, V., Sydorenko, T., & Daurio, P. (2020). Using spoken dialogue technology for L2 speaking practice: What do teachers think? *Computer Assisted Language Learning*, 1–24. <https://doi.org/10.1080/09588221.2020.1774904>
- Timpe-Laughlin, V., Sydorenko, T., & Dombi, J. (2022). Human versus machine: Investigating L2 learner output in face-to-face versus fully automated role-plays. *Computer Assisted Language Learning*, 1–30. <https://doi.org/10.1080/09588221.2022.2032184>
- Tran, C. T., Zhang, M., Andreae, P., Xue, B., & Bui, L. T. (2018). Improving performance of classification on incomplete data using feature selection and clustering. *Applied Soft Computing*, 73, 848–861. <https://doi.org/10.1016/j.asoc.2018.09.026>
- Tricco, A. C., Lillie, E., Zarin, W., O'Brien, K. K., Colquhoun, H., Levac, D., et al. (2018). PRISMA extension for scoping reviews (PRISMA-ScR): Checklist and explanation. *Annals of Internal Medicine*, 169(7), 467–473.
- Tripathi, M., Ray, N., Sinha, A., & Ahmed, F. (2021). New media and language learning: Opening global doors to linguistic development. *International Journal of English Learning & Teaching Skills*, 3(4), 2507–2566. <https://doi.org/10.15864/ijelts.3407>
- Tsay, C. H. H., Kofinas, A. K., Trivedi, S. K., & Yang, Y. (2020). Overcoming the novelty effect in online gamified learning systems: An empirical evaluation of student engagement and performance. *Journal of Computer Assisted Learning*, 36(2), 128–146.
- Ulum, Ö. G. (2020). A critical deconstruction of computer-based test application in Turkish State University. *Education and Information Technologies*, 25(6), 4883–4896.
- Underwood, J. H. (1984). *Linguistics, computers, and the language teacher. A communicative approach*. ERIC.
- Vollstedt, M., & Rezat, S. (2019a). An introduction to grounded theory with a special focus on axial coding and the coding paradigm. *Compendium for Early Career Researchers in Mathematics Education*, 13(1), 81–100. <https://doi.org/10.1007/978-3-030-15636-7>
- Vollstedt, M., & Rezat, S. (2019b). An introduction to grounded theory with a special focus on axial coding and the coding paradigm. *Compendium for Early Career Researchers in Mathematics Education*, 13, 81–100. [https://doi.org/10.1007/978-3-030-15636-7\\_4](https://doi.org/10.1007/978-3-030-15636-7_4)
- Wahyuningsih, S., & Afandi, M. (2020). Investigating English speaking problems: Implications for speaking curriculum development in Indonesia. *European Journal of Educational Research*, 9(3), 967–977. <https://doi.org/10.12973/eujer.9.3.967>
- Wang, Y., Feng, X., Guo, J., Gong, S., Wu, Y., & Wang, J. (2022). Benefits of affective pedagogical agents in multimedia instruction. *Frontiers in Psychology*, 12, 6639. <https://doi.org/10.3389/fpsyg.2021.797236>
- Weizenbaum, J. (1976). *Computer power and human reason: From judgment to calculation*. Wolfsinkel, J. F., Furtmueller, E., & Wilderom, C. P. (2013). Using grounded theory as a method for rigorously reviewing literature. *European Journal of Information Systems*, 22(1), 45–55. <https://doi.org/10.1057/ejis.2011.51>
- Wollny, S., Schneider, J., Di Mitri, D., Weidlich, J., Rittberger, M., & Drachsler, H. (2021). Are we there yet? A systematic literature review on chatbots in education. *Frontiers in artificial intelligence*, 4.
- Wong, G. K., Ma, X., Dillenbourg, P., & Huan, J. (2020). Broadening artificial intelligence education in K-12: Where to start? *ACM Inroads*, 11(1), 20–29. <https://doi.org/10.1145/3381884>
- Wright, M. (2022). *Building connections in 2021–22*. <https://www.britishcouncil.org/about-us/how-we-work/review-2021-22>
- Wu, C. H., Lin, H.-C. K., Wang, T.-H., Huang, T.-H., & Huang, Y.-M. (2022). Affective mobile language tutoring system for supporting language learning. *Frontiers in Psychology*, 13, Article 833327.
- Wu, Y., Rough, D., Bleakley, A., Edwards, J., Cooney, O., Doyle, P. R., et al. (2020). See what I'm saying? Comparing intelligent personal assistant use for native and non-native language speakers. In *Proceedings of the 22nd international conference on human-computer interaction with mobile devices and services* (pp. 1–9). <https://doi.org/10.1145/3379503.3403563>
- Wu, Shirkey, G., Celik, I., Shao, C., & Chen, J. (2022). A review on the adoption of AI, BC, and IoT in sustainability research. *Sustainability*, 14(13), 7851. <https://doi.org/10.3390/su14137851>
- Xia, P. (2020). *Application scenario of artificial intelligence technology in higher education* international conference on applications and techniques in cyber intelligence ATCI 2019. In *Applications and techniques in cyber intelligence* (Vol. 7).
- Xie, Q. (2017). Investigating the target language usage in and outside business English classrooms for non-English major undergraduates at a Chinese university. *Cogent Education*, 4(1), Article 1415629. <https://doi.org/10.1080/2331186X.2017.1415629>
- Xie, T., Liu, R., Chen, Y., & Liu, G. (2021). Moca: A motivational online conversational agent for improving student engagement in collaborative learning. *IEEE Transactions on learning technologies*, 14(5), 653–664.
- Yang, A. C., Chen, I. Y., Flanagan, B., & Ogata, H. (2022). How students' self-assessment behavior affects their online learning performance. *Computers & Education: Artificial Intelligence*, 3, Article 100058. <https://doi.org/10.1016/j.caeai.2022.100058>
- Yin, Q., & Satar, M. (2020). English as a foreign language learner interactions with chatbots: Negotiation for meaning. *International Online Journal of Education and Teaching*, 7(2), 390–410. <https://iojet.org/index.php/IOJET/article/view/707>
- Young, R. F. (2011). Interactional competence in language learning, teaching, and testing. In *Handbook of research in second language teaching and learning* (pp. 426–443). Routledge.
- Yu, Y., & Chen, Y. (2018). Design and development of high school artificial intelligence textbook based on computational thinking. *Open Access Library Journal*, 5(9), 1–15. <https://doi.org/10.4236/oalib.1104898>
- Zhai, C., & Wibowo, S. (2022). A systematic review on cross-culture, humor and empathy dimensions in conversational chatbots: The case of second language acquisition. *Heliyon*, Article e12056. <https://doi.org/10.1016/j.heliyon.2022.e12056>

- Zhai, C., Wibowo, S., & Cowling, M. (2022). *Work-in-Progress—embedding cross-cultural humorous and empathetic functions to facilitate language acquisition*. 2022 8th International Conference of the Immersive Learning Research Network (ILRN). <https://doi.org/10.1016/j.heliyon.2022.e12056>
- Zhang, J., & Jing, Y. (2022). Application of artificial intelligence technology in cross-cultural communication of intangible cultural Heritage. *Mathematical Problems in Engineering*. <https://doi.org/10.1155/2022/6563114>, 2022.
- Zhang, J., Mandl, H., & Wang, E. (2011). The effect of vertical–horizontal individualism–collectivism on acculturation and the moderating role of gender. *International Journal of Intercultural Relations*, 35(1), 124–134.
- Zhang, Y., Zhao, M., Ouyang, X., & Zhao, Y. (2016). Study on knowledge engineering education technique. In *The 2016 international conference on education* (pp. 27–31). E-learning and Management Technology. <https://doi.org/10.2991/iceemt-16.2016.6>.
- Zou, B., Liviero, S., Hao, M., & Wei, C. (2020). Artificial intelligence technology for EAP speaking skills: Student perceptions of opportunities and challenges. In *Technology and the psychology of second language learners and users* (pp. 433–463). Springer.
- Zrekat, Y., & Al-Sohbani, Y. (2022). Arab EFL University learners' perceptions of the factors hindering them to speak English fluently. *Journal of Language and Linguistic Studies*, 18(1), 775–790. <https://doi.org/10.52462/jlls.219>