AI Tools to Support English Language Learning in K-12 Classrooms

THESIS

Presented in Partial Fulfillment of the Requirements for the Degree Master of Arts in the Graduate School of The Ohio State University

By

Mauricio Antonio Dos Santos

Graduate Program in Education Teaching and Learning

The Ohio State University

2025

Thesis Committee:

Professor Youngjoo Yi, Advisor

Professor Ian Wilkinson

Abstract

Artificial intelligence has influenced the field of English language learning in different aspects, and several applications have been developed to support English Language Learners (ELLs) and teachers. Due to their advancements and popularity among teachers, students, and administrators, using AI tools for learning languages has been a trending topic. Thus, scholars have published works about how to use AI tools in different settings to support the input, output, and interaction of humans and machines in the process of Second Language Acquisition (SLA). Studies on the use of AI tools for teaching and learning English in higher education have highlighted both the positive impact on language learning and concerns regarding data privacy and academic integrity. However, there is still a research gap specifically addressing the use of AI tools in ESL/EFL K-12 classrooms. To fill this gap, further studies on the use of AI tools in ESL/EFL K-12 contexts are necessary to understand how these tools are used to support teachers and learners who are already using AI tools inside and outside the classroom but lack insight into the extent of support these tools can provide. This narrative research synthesis informed by the input, output, and interaction hypothesis by Krashen, Swain, and Long (1985) and taxonomy of AI tools in education by Holmes and Tuomi (2022) aimed to investigate what AI tools are used in the context of ESL/EFL K-12 classrooms, the English skills they are used to supporting, the perceptions of teachers and students about effectiveness, challenges, and limitations of these tools, the topics regarding the use of AI tools and the recommendations regarding the use of AI tools. Five research questions were used to guide this synthesis. Twenty-three studies conducted between 2022 and 2024 were selected for this review and analyzed using a coding sheet aligned with the five research

questions. The analysis of the studies showed that AI tools were more used in Asian countries in the context of EFL. Chatbots and AI Personal Assistants were the most used AI tools, and they supported speaking reading, writing, and listening skills through vocabulary acquisition, pronunciation, and corrective feedback and recast. Most AI tools focused on speaking skills and promoted the development of a student-centered approach inside and outside the classroom. In addition, the data showed how the extent of support given by AI tools might be influenced by the English proficiency level of the students, their level of anxiety, and the environment in which they are immersed. Recommendations for future research about the concerns with data privacy and academic honesty were suggested as well as a more representative sample for future studies including different contexts and specific needs of students according to their level, needs, and motivation.

Acknowledgments

I am deeply grateful to my advisor, Dr. Youngjoo Yi, for her unwavering guidance and support. She is one of the most knowledgeable, kind, and supportive professors I have had the privilege of working with at The Ohio State University. Since the beginning of my MA program, she has guided me in all aspects of my academic journey.

I am deeply grateful to my committee member, Dr. Ian Wilkinson. It has been an honor to have him as both my professor and committee member. His guidance and insights have greatly contributed to this work, as well as to my personal and professional growth."

I am deeply thankful for my family. To my wife Kelly, for her love and unconditional support; my daughter Mariana, whose beautiful drawings always brightened my days; and my son Davi, for his kindness whenever I needed it. Coming to the USA together made this journey one of the best experiences of our lives. I also thank my parents, Ademar and Maria, and all my brothers and sisters for their love and encouragement.

I thank my colleagues from BRASA and Buckeye Brighters for enriching my academic life. I also thank my professors, including Dr. Cristiane Rodrigues and Dr. Michel Macedo Marques for their support throughout my journey. To all my friends and students —thank you for being part of this experience.

I sincerely thank SEDUC-CE, CAPES, and Fulbright Brazil, especially Dr. Luis Loureiro, Dr. Alexandre Prestes, and Carolina Martins, and Liduina Patricio for their support since the beginning of this academic program.

Above all, I thank God for the gift of life.

Vita

August 2005……(Licenciatura) Language Arts, Universidade Regional do Cariri

2011-2014……(Graduate Certificate) English Language Teaching, Universidade Regional do Cariri

2015-2017.....English professor, Universidade Regional do Cariri

2015.....English Teacher, EEMTI Tiradentes

2023 to present .....M.A. Education of Teaching and Learning, The Ohio State University

Fields of Study

Major Field: Education Teaching & Learning

Table of Contents

List of Tables

List of Figures

Chapter 1: AI Tools to Support English Language Learning in K-12 Classrooms

In the field of Second Language Acquisition (SLA), different methods and approaches have influenced the way languages are learned, (Larsen-Freeman & Anderson, 2011; Lightbown & Spada, 2013; Ortega, 2009). Recently, several types of innovative technologies have been applied to and influenced the field of English language learning (ELL) (John, 2018; Ortega, 2017). Among the emergent technologies available in our society, Artificial Intelligence (AI) tools have become more popular and accessible for users of technological devices, such as smartphones and computers. Consequently, current language learning standards in the USA and other countries have included the use and understanding of innovative technologies as a priority for global development and positioning. For instance, the American Council on the Teaching of Foreign Languages (ACTFL) in a position statement from May 19, 2017, highlighted the importance of technology and recommended its use by language educators as a tool to support language learning.

This same view, including some concerns about its impact on humans, is shared in the book *AI and education: Guidance for policy-makers* (UNESCO, 2021). This book defined what artificial intelligence is in the context of education and its role in the development of teaching and learning through the progress of Sustainable Development Goal 4 (SDG 4) which aims to promote quality education. They also released a report entitled *Technology in education: A tool on whose terms?* (UNESCO, 2023) which included additional discussions on how technology should be regulated in society. Finally, the International Society for Technology (ISTE) also released material with standards and recommended ways for the use of technology in education to support teachers, students, and administrators.

Due to AI tools advancements and accessibility in distinct levels of the educational system. Teachers, students, and administrators are already using them inside and outside the classroom for diverse activities, including learning languages. Consequently, this use does not consider further challenges and implications about the impact of AI tools on educational settings. This lack of research has raised global discussions and several questions about the best ways to implement AI tools in learning and assessment settings to measure the efficacy of applications, such as intelligent tutoring systems (ITS), automated writing evaluation (AWE), AI-supported language learning, and their support for fostering the human rights and equitable education (UNESCO, 2021).

The theoretical framework of this work is based on the works of Krashen, Swain, and Long (1985) and their respective hypothesis on comprehensible input, output, and interaction (Gass & Mackey, 2014; Lightbown & Spada, 2013; Long, 2017; Ortega, 2009). These hypotheses play a crucial role in building the understanding of how languages are learned. It is important to point out that studies about AI, and English language learning have shared some connections with SLA. For instance, Hsu et al. (2024) stated that when young learners interact with AI image recognition technologies for vocabulary acquisition, they engage with text-based, non-text, and object input, in addition to output, and reflection. Ericsson et al. (2023) mentioned that when students use virtual human chatbots, they negotiate meaning what is connected to Long (1985, 1996). In addition, they can interact with comprehensible input, output, and feedback, which aligns with Chen et al. (2022) who noted that speech recognition technology provides corrective feedback to promote self-repaired output. Jeon (2023, 2024) discusses the interaction between

students and chatbots, suggesting how the use of chatbots for dynamic assessment can promote interactive glossing, which is important for meaning-focused input according to Nation (2001).

Input, interaction, feedback, and output, are important for fostering an interactional approach when students use AI tools. Several authors (i.e., Tai & Chen, 2023, 2024; Wang et al., 2024; Wang et al., 2022; Woo et al., 2024) emphasize the importance of humans learning languages through input, output, and interaction, paralleling how AI tools receive the input from humans, process it, and generate output. This parallels the input-process-output (IPO) model, which is present in various disciplines, including SLA as proposed by Krashen, Swain, and Long (1985). This connection between SLA theories and AI tools is evident, as topics such as Natural Language Processing (NLP), Spoken Language Identification (SLI), Machine Learning (ML), and Speech Recognition (SR) share common ground in both fields.

The connection between SLA and AI shows that reflecting on what happened in the past while looking at the present might influence the development of quality research in both fields. These ideas can be used to guide teachers, students, and policymakers on how AI tools in educational settings might help them to see specific needs they might have about English language learning. Even before the 1950s, there were already influential disruptive ideas that contributed to the fields of AI and education, and they are still relevant. However, for this work, I will focus on the development of AI tools after the 1950s and their application in educational settings, specifically English language learning in ESL/EFL classrooms.

AI Tools History

When Turing (1950) asked the question “Can machines think?” (p. 433) and started to explain the definition of the terms “machine” and “think”, that led him to ask other questions and

consequently to develop the *imitation game* "Turing Test", a test to compare the intelligent behavior of a machine to the behavior of a human. His work has influenced the field of AI and technology in distinct aspects. Then, in 1956 during the Dartmouth Summer Research Project, John McCarthy coined the term Artificial Intelligence (Radanliev, 2024). Apart from the field of computer science, another scholar whose work also contributed to the field, specifically the part of Natural Language Processing (NLP) was Noam Chomsky. In 1957 he published "Syntactic Structures", which became popular in the fields of linguistics and computer science as well. Consequently, other famous scholars such as Charles Babbage, Dennis Richie, and Tim Paterson followed the development of Artificial intelligence through the years by giving their contributions.

Among the AI tools developed through the years, chatbots have become the most popular ones. The first documented Chatbot was a type of Rogerian psychotherapist chatbot called *ELIZA* developed in the mid-1960s by Professor Joseph Weizenbaum at MIT. *ELIZA* used Natural Language Processing (NLP) to make conversation between a human and a machine possible (Weizenbaum, 1966). Thus, the development of AI tools after *ELIZA* kept increasing. Finally, in 2015, Sam Altman and Elon Musk co-founded a company called Open AI. Then in November 2022, Open AI launched ChatGPT, an application based on Generative Pre-Trained Transformer (GPT-3) (OpenAI, 2023). ChatGPT is a Large Language Model (LLM) more advanced than *ELIZA* with the ability to adapt the user's input creating a human-like conversational dialogue (Huang et al., 2022).

According to Radanliev (2024), ChatGPT is a superior AI model leading to a new paradigm shift in the field of AI. This shift was followed by other companies that developed their

own AI tools, such as Google (Gemini), Microsoft (Copilot), Amazon (Alexa), and Apple (Siri). These advancements also brought to the users some specific functionalities for the AI tools, such as speech recognition, text recognition, image recognition, translation, and personal assistance. These AI tools have opened several possibilities for the integration and improvement of other technological tools in the field of education. Among them, learning language assistants like *Duolingo,* which is boosted by AI models like GPT (Bicknell et al., 2023), *NotebookLM*, a new tool released by *Google* capable of summarizing sources and generating an audio interactive conversation about it, and *Turnitin,* a plagiarism detection and prevention service that uses Natural Language Processing (NLP) to detect AI generated texts.

AI tools can perform several activities in different contexts, and this has also caused some concerns about it replacing humans. UNESCO (2021) suggests that all the implementation of AI tools should focus on humans. Barrot (2023) also highlighted these concerns. For him, AI tools are and will become part of the process of teaching and learning, in his words although ChatGPT and other AI tools can act like humans they cannot replace humans because they cannot help individuals to use their voices, identities, and distinctiveness in writing. In addition, he emphasizes that AI tools do not have emotional depth and life. He concludes his thoughts by saying that AI tools cannot replace humans. While he mentioned that AI tools cannot replace teachers, An et al. (2023) showed that schools that use AI tools still need teachers.

AI Tools in Education

The development of AI tools in fields like science, engineering, math, and education has increased globally. For this reason, some countries have tried to implement policies and regulations about the use of AI tools in educational settings. On September 15, 2017, China's

State Council released a document entitled “Next Generation Artificial Intelligence Development Plan” to guide the new policies in education to support the implementation of AI tools to create what they called “intelligent education” and on July 13, 2024, they released another document called “Generative AI Regulation”. Other countries have released their own AI policies as well. In Brazil, for example, bill No.2,338/2023 was approved in December 2024 to regulate the use of AI, including in education, while in Europe, AI is regulated by the EU AI Act released in 2023.

Although regulations and policies were important to guide the legal aspect of using AI tools in education, it was also necessary to investigate how these tools could be applied in educational settings. Holmes and Tuomi (2022) discussed their expectations and the impact of Artificial intelligence in education (AIED). Their work focused on how innovative AI tools could merge with other existing technologies, contribute to broadening perspectives with their use in education, and how they could support educators and foster the skills of learners. They also included discussions on how our society should consider the functions of education, such as socialization, social integration, social placement, and social and cultural innovation to have better results in applying these tools. Similarly, Burkhard (2022) highlighted the importance of instructors conducting the students in a more preventive approach to reduce the risks or misuse of AI tools. He emphasized how discipline influences the use of these tools in a way that both educators and students can understand their functionalities and how to use them. Finally, Yan (2023) pointed out that there are concerns about using AI tools, specifically ChatGPT, but he also highlighted the benefits of using these tools to overcome the challenges (e.g., academic integrity, plagiarism).

AI tools inside and outside of ESL/EFL classrooms

Currently, the terms ESL, EFL, and L2 are interchangeably related to Second Language Acquisition (SLA). For this work, I will consider ESL in the context of countries (e.g., USA, Canada, Ireland), where people (e.g. immigrants) learn English as a second language and EFL in the context of countries where English is not the official or dominant language, and people learn it as a foreign language (e.g., Brazil, China, Japan). However, I might present the terms in other circumstances due to the definitions given by other authors in their studies. This definition of the terms ESL/EFL is an important way to understand the dynamics of the environment and clarify the geographical and level of interaction the learners have with their target language or L2. For example, the context of learning English for individuals living in or out of the country might change significantly, when in the country some people learn English as a foreign language, but when they move to another country where English is the official language, they learn English as a second language. This knowledge about the process of learning languages is dynamic and beyond the context of the country where the learner is. There is also the context of learning inside and outside the classroom (Benson, 2011), sometimes described as formal and informal settings (Bahrani et al. 2014; Krashen, 1981; Long 1996; Ortega 2009; Swain, 1985). In this work, the terms inside the classroom (formal) will be defined as any setting where the teachers plan and conduct formal instruction for students. In this case, it can be in person, online, synchronous, or asynchronous. Outside the classroom (informal) is everything students might use voluntarily or not to cover their specific needs in learning any language they are interested in.

In this work, I will not focus specifically on the process of AI-mediated informal digital learning of English (AI-IDLE), which refers to learning English outside structured classrooms by

using AI tools (Liu et al., 2024) but discuss the possibility of using AI tools in formal and informal environments according to the findings from the research available. This is important to consider because when AI tools are applied in different contexts, they might influence the extent of the impact on the process of learning a language. For example, *Duolingo* is an application that students use to learn languages outside the classroom. Duo, the owl (the mascot of Duolingo) sends notifications to the learners about what they need to do to keep their learning progress (Bicknell et al., 2023), but Duolingo offers the option for teachers include it in the curriculum and assign activities for the students. Diliberti et al. (2024) showed how popular adaptive learning systems (e.g., *Khan Academy)* are used by teachers in the USA to support instruction in the classroom, but it is also used by students who want to study by themselves outside the classroom.

In countries where English is taught as a foreign language in a large class size, some students do not have the opportunity to practice their speaking skills and to receive personalized feedback from their instructors. Consequently, they look for support out-of-school and engage with technology as one alternative to cover their specific needs. Other situations might also contribute to this distancing that students face inside and outside the classroom, such as the method the teachers use, for example, the grammar-translation method was used when reading was the domain teachers were more focused on. It made the classroom more teacher-centered, and the students were assessed only by one specific skill. There was little interaction between teachers and students.

Now, the standards for English learning in several countries have changed and they encourage teachers to focus on the four English skills (reading, writing, speaking, listening). In Brazil, for example, The Brazilian National Common Base (BNCC) is a curriculum that aims for

Brazilians to develop all domains of English language learning. However, teachers and students are looking for alternatives like AI tools to practice the four skills when they do not receive support in the classroom. In general, AI tools have been one option to cover the skills students do not have the chance to improve or when the teachers do not have time to give feedback to them. As an instructor myself, I have taught classrooms with 45 high school students in Brazil, and one of the biggest challenges was to give feedback to them because I did not have enough time to do it.

According to Yeh (2024), traditional pedagogical approaches are limited and require a more authentic learning environment by integrating AI tools to support students and teachers in the context of EFL education. The same view is shared by Holmes and Tuomi (2022) who presented a taxonomy of AIED systems that was divided into three parts "(1) student-focused, (2) teacher-focused, and (3) institution-focused AIED" (p. 550). Even though this taxonomy was created as a reference for research about the context of AI in education (AIED), it can be applied to research in the context of English language learning. AI tools are used by language instructors and English learners (ELs), as Diliberti et al. (2024) mentioned, in the context of K-12 education most users of AI tools were language arts teachers and students.

Regarding the most recent AI tools, Konyrova (2024) conducted a survey to investigate the impact of AI in the ESL context and found out that seventy percent of ESL classrooms use some type of AI, in which sixty-five percent are language learning apps, fifty-five percent are AI-based grammar and writing checkers, and forty percent are driven pronunciation correction tools. Among the AI tools in EFL teaching and learning settings, the most used according to Alshumaimeri and Alshememry (2024) are Machine Learning (ML), Deep Learning (DL), and

Computer Vision (CV). Their systematic review included studies published between 2007 and 2022 only in the context of EFL learning skills in general, they did not focus on a specific group of students, just EFL students. Most authors agreed that AI tools used in the context of ESL/EFL presented benefits and challenges related to their use, but most studies were focused on higher education (e.g., Mustroph & Steinbock, 2024). For this reason, it is important to understand the relationship between AI tools in the context of ESL/EFL in higher education and ESL/EFL K-12 classrooms, as well as whether they demonstrate the same level of efficacy.

Reflecting on the use of AI Tools in K-12 Education in the USA, a Research AND Development (RAND) report found that 27 percent of language arts teachers used AI tools in their work while only 11 percent of STEM teachers did the same (Diliberti et al., 2024). This percentage of users highlights that AI tools are already been used in the classrooms, even before their efficacy is tested, the survey indicated that virtual learning platforms, adaptive learning systems, and chatbots were among the most common. Some teachers mentioned they were using AI tools in the classroom weekly. This report included a sample of 1,020 K-12 teachers to understand how they used AI tools, 231 districts, and 11 district leaders to understand how they were implementing the policies to support the teachers. Their findings indicated that 60 percent of districts were planning to train their teachers about how to use AI tools, among them, the most popular tools listed were, virtual learning platforms (e.g. *Google Classroom),* Adaptive learning systems (e.g., *Khan Academy)*, and Chatbots (e.g. *ChatGPT, Google Bard)* (Diliberti et al., 2024). Other AI tools included were Machine Translation (MT) tools, text-to-speech, speech-to-text, and writing assistants (Jiang et al., 2023).

Benefits and Challenges of using AI tools

With respect to the benefits, AI tools like chatbots are used to support the students' reading and writing skills (e.g., Barrot 2023; Burkhard, 2022; Yan, 2023), while others are available to foster all language skills such as reading, writing, speaking, and reading (e.g., Holmes & Tuomi, 2022; Huang et al., 2022; Konyrova, 2024). According to Alhalangy and AbdAlgane (2023), AI has a positive impact on the ESL/EFL context because it facilitates pedagogical and learning procedures. Similarly, Shafiee Rad & Roohani (2024) found in their empirical study that students using AI tools in individualized learning had more positive attitudes and improved their pronunciation when compared to their peers. These findings about personalized learning show how students can build their knowledge (Kim et al., 2023), and how meaningful a more student-centered approach can be used to foster the performance of students in the context of L2 language learning (Lightbown & Spada, 2013; Ortega, 2009). Another finding mentioned by Kasneci et al. (2023) is how AI tools like ChatGPT can support students with disabilities such as visual impairment when they use speech-to-text or text-to-speech solutions in their language learning journey.

With respect to the challenges, Alshumaimeri and Alshememry (2024) stated that in the context of ESL/EFL, data safety, privacy, access to AI tools, lack of spontaneous improvisation, inventiveness, and shared understanding are the most common issues. They also stated that chatbots could be inconsistent, confusing, and prone, and Machine Translation (MT) could be disruptive, unsuccessful, and contribute to the violation of the learner's integrity. Other authors agreed with them about the issues related to academic integrity (Burkhard, 2022; Holmes & Tuomi, 2022; Yan, 2023). In the context of academic writing Barrot (2023) highlighted the

challenge ChatGPT has in capturing the writing voices and identities of the students while they use it. Finally, Konyrova (2024) stated the challenge of how to balance humans and machines in the integration of AI tools in ESL classrooms, she mentioned how training is important to support educators in this aspect.

Another challenge about the use of AI tools in the classroom is the extent of support teachers and students receive from policymakers. Diliberti et al., (2024) mentioned that only five percent of the districts had adopted specific policies for students about the use of AI tools. Without the policies, students and teachers might have questions about the use, and these questions are not answered, so this increases the level of uncertainty. So, the focus of the districts is more on teachers' training, so they can learn and focus on the creation of a more student-centered environment based on the suggestions of several standards recently released. Diliberti et al., (2024) indicated that teachers in different districts intend to use AI tools in the future and policymakers intend to implement the regulations. This requires more studies, partnerships, support, and good intentions from all stakeholders to make this transition to intelligent systems as smooth as possible, considering several concerns highlighted by UNESCO (2021).

Problem Statement

Despite the increasing amount of research on the use of Artificial Intelligence at different levels and for different domains of teaching and learning English (Alshumaimeri & Alshememry, 2024; Konyrova, 2024), there is a gap in the context of using AI in ESL/EFL K–12 education and it should be more explored (Lo et al., 2024). Most current studies about AI tools in the ESL/EFL settings are related to students at colleges and universities, as well as independent adult language learners. It might happen due to difficulties related to regulations to implement the use of these

tools in K-12 settings, the lack of teacher training, the perception that K-12 students are immature, and the lack of a more student-centered environment for ELLs. However, by considering the new generations and their development of 21st-century skills. Several national standards have considered the integration of technology and innovation to support the development of those skills. To do so, it is important to understand the effectiveness of the AI tools available for English language learning in K-12 classrooms and how they have been implemented to support the students inside and outside of the classroom.

Through this narrative research synthesis, I aim to review current empirical studies in terms of the attitudes, utilization, contribution, and outcomes of AI tools in supporting English Language Learning in ESL/EFL K-12 classrooms in order to discover the existing gaps in research. To do so, I intend to identify what AI tools are more currently used, how they have been used to support English language learning in ESL/EFL K-12 classrooms, what the perceptions of teachers and students about their use are, and to point out the main findings and recommendations from the studies selected for the review. In addition, I intend to explore whether AI tools can also be used to support students when they are outside the classroom. Ortega, (2009) and Yi, (2021) highlighted how the students' perceptions, life experiences, connections, and background knowledge might influence the way they interact with the languages. This influence of formal and informal settings might indicate the reasons the students need extra support to achieve their goals. AI tools are perceived as machines, and they are used in several environments inside and outside the classroom. An understanding of how they have been used in different contexts might be useful for guiding teachers and students in the future.

I expect this work will be a source of information for teachers, students, and institutions, as well as a contribution to the field with a comprehensive understanding of the most recent AI tools used for supporting the learning of the English language in ESL/EFL K-12 classrooms. The following research questions (RQs) guided me during the investigation of the empirical studies selected:

RQ1. What AI tools are used to support the learning of the English language inside and outside the ESL/EFL K-12 classroom?

RQ2. What English language skills (listening, speaking, reading, and writing) are taught with the help or support of AI tools in ESL/EFL K-12 classroom?

RQ3. How do teachers and students perceive the effectiveness, challenges, and limitations of AI tools in ESL/EFL K-12 classroom settings?

RQ4. What are topics regarding the use of AI tools in ESL/EFL K-12 from the studies reviewed?

RQ 5. What are the recommendations regarding the use of AI tools in ESL/EFL K-12 from the studies reviewed?

To answer these questions, I designed this study around the input, interaction, and output hypothesis proposed by Krashen, Swain, and Long (1985), the taxonomy designed by Holmes and Tuomi (2022), and recommendations given by international organizations from different countries about using AI tools for learning languages. Other previous studies and guidelines about English language learning and Artificial intelligence were useful in building an understanding of the arguments about the use of AI tools and ESL/EFL K-12 classrooms. These questions were

revisited in this narrative research synthesis every time I needed to interpret the information from the empirical studies selected. I have divided this work into five chapters and different sections to immerse the reader in a sequence of innovative, challenging, and promising ideas from studies conducted in different countries, with different perspectives, and different approaches on how AI tools can support English language learning in ESL/EFL K-12 classrooms.

This is the road map to guide the reader during this work: The first chapter of this work is the introduction of the most popular topics about AI tools to give a background about the uses and advancements in education and English language learning. The second chapter is the method section where I describe the steps taken to answer the five research questions. In the third chapter, I highlighted the main findings of the studies reviewed and presented them according to their level of impact on the field. The fourth chapter is the discussion, which I divided into five sections; each section discusses each one of the research questions. The fifth and last chapter is the conclusion about the work and recommendations for future research to give a reader a perspective of what is happening and about what is to come in this marvelous world of AI tools and language learning.

Chapter 2: Method

This work is a narrative research synthesis guided by the processes of research synthesis provided by Shanahan (2000), and the procedures of synthesizing research in English teaching and learning by Norris and Ortega (2006), as well as Ortega (2015). It also included the four elements of a narrative synthesis described by Rodgers et al. (2009), (1) developing a theory, (2) developing a preliminary synthesis, (3) exploring relationships within and between the studies, and (4) assessing robustness, conclusions, and recommendations. The choice for a more systematic narrative synthesis looks relevant in the context of the continuous development of recent technologies. This work is qualitative and might address diverse types of data from qualitative and quantitative research. It also contributes to the understanding of the functionalities and applications of AI tools.

Narrative research synthesis is a good option because the topics of studies about AI tools are in the process of development, and they normally bring aspects of perceptions, behaviors, effectiveness, and impact. The innovative uses of AI tools in different K-12 settings present several variables that are measured quantitatively and qualitatively and might present better results if analyzed together. These studies might present a diversity of information that requires an approach where some aspects that go beyond statistics can contribute to the field in general by giving a distinct perspective of what is currently happening or might happen in the future. In addition, it is important to build a concise timeline of their implementation and efficacy through the years and places to see how the use of these tools has changed and influenced the learning of languages.

Literature Search

To identify the studies for this synthesis, I conducted a systematic electronic search using as inclusion criteria only peer-reviewed academic journals from two databases: ERIC and PsycINFO (APA). Based on the research questions, I used Boolean operators and included the following search terms: “Artificial Intelligence” OR “AI” OR “A.I.” AND “English as a second language” OR “ESL” OR “English as a foreign language” OR “EFL” OR “English language learning.” I found 531 studies between 1980-2024. filtering only peer-reviewed and academic journals, I found 395 studies in ERIC and 92 studies in PsycINFO. As a limiter, I included only studies between 2022 and 2024, and 330 were retrieved. After excluding 6 duplicated studies, I conducted a manual final screening through reading titles and abstracts, I included only empirical studies in the context of ESL/EFL K-12, resulting in 36 empirical studies. A final screening after reading the titles, abstracts, and in some studies full text resulted in 23 empirical studies about AI tools in the context of ESL/EFL K-12 classrooms.

Inclusion and Exclusion Criteria

I have chosen studies published from 2022 to 2024 and removed the studies that did not cover these criteria, the date when the studies were published worked as exclusion criteria for 113 papers in this work. Three years is not a common standard for research synthesis. However, previous research synthesis about AI tools in the EFL context showed that most studies were published after the release of ChatGPT in 2022 (e.g., Alshumaimeri & Alshememry, 2024). Additionally, according to OpenAI (2023), the first GPT model was released in June 2018, then GPT-2 in February 2019, GPT-3 in June 2020, GPT-3.5 in November 2022, which was the base

for the release of ChatGPT, and GPT-4 in 2023. This fast development of ChatGPT and other AI tools showed how short the interval between the launching of new tools and updates is.

I developed both inclusion and exclusion criteria for this research. The inclusion criteria include as follows:

(1) peer-reviewed academic articles written in English language

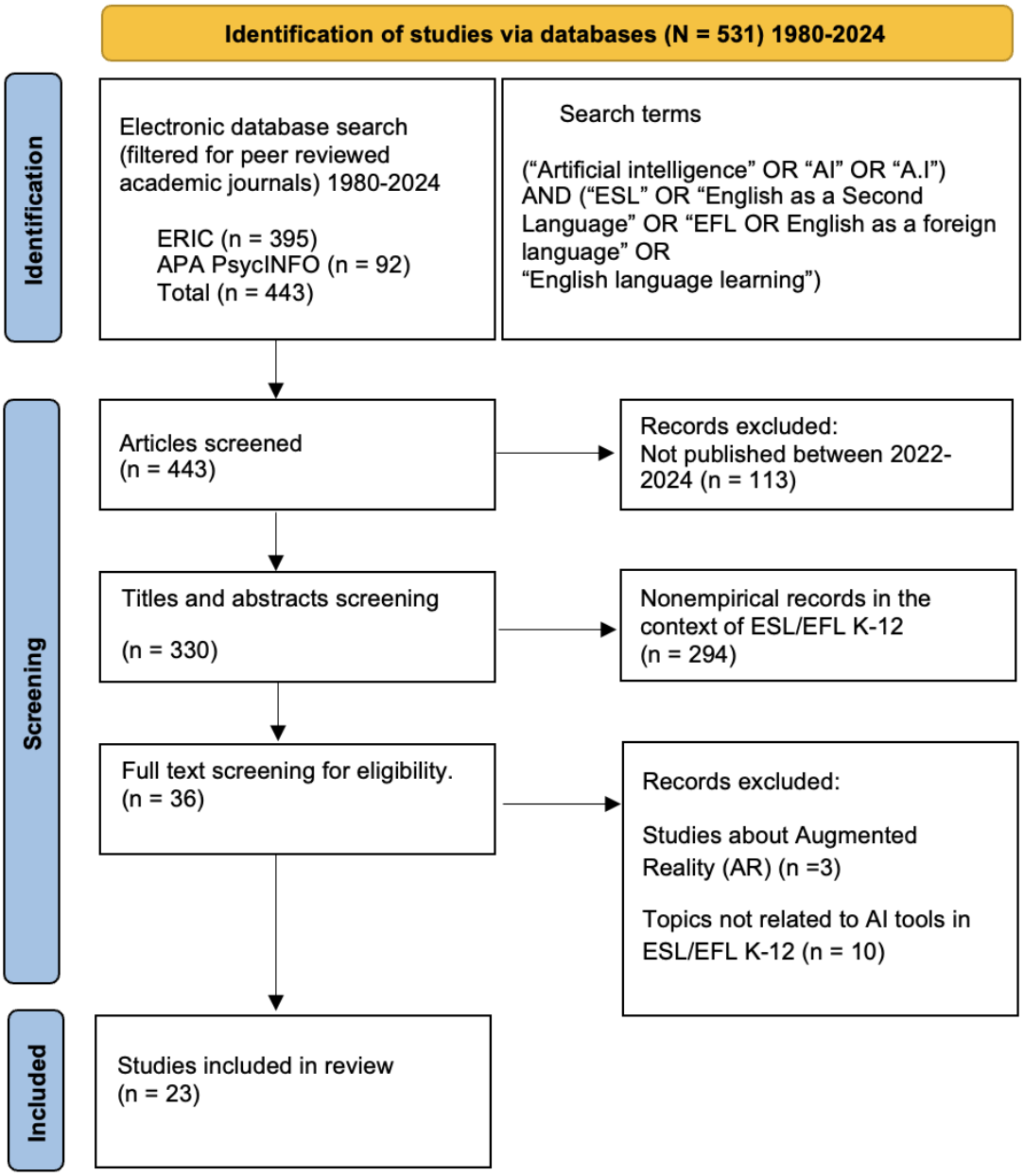
(2) empirical studies that could be qualitative, quantitative, and/or mixed methods

(3) studies that examined the use and/or implementation of AI tools as well as the perceptions of teachers and students in ESL/EFL K-12 classroom

(4) studies published from 2022 to 2024

As exclusion criteria, I excluded (1) studies conducted beyond K-12 contexts (e.g., higher education), (2) studies about the design of AI tools without any type of experiment or implementation with students, (3) studies examined the use of AI tools beyond English language learning and teaching contexts, (4) studies about AI tools like Augmented Reality (AR). Studies about Augmented Reality (AR) were not included in this paper because according to Holmes and Tuomi (2022), AR are not AI technology themselves, but they are combined with "AI machine learning, image recognition, and material language processing and are increasingly being used in educational settings" (p. 552). Thus, Augmented Reality research was not included in this review.

Figure 1 *Prisma chart of the search*



Note. This flow chart describes the process of identification, screening, and inclusion of the studies.

Data Analysis

I used both deductive and inductive approaches to analyze the data in this study. At the very first stage of data analysis, I used a deductive coding approach based on the five research questions and created a coding sheet divided into five predefined categories; (1) General information (title, reference, year, source, study, country/state, setting (ESL/EFL classrooms), aim of the study, research questions, methodology, participants (students/teachers), English/education level of the students, (2) AI tools used in ESL/EFL K-12 (names of AI tools, types of AI tools, frequency of use) (3)) specific English language skills (reading, writing, listening, speaking) taught by AI tools. (4) Perceptions of teachers/students about AI tools (effectiveness, challenges, bias/limitations), and (5) Findings and recommendations (effectiveness, challenges, bias/limitations, effect size, recommendations, study design, sample size, data collection). As such, I coded the data in a deductive manner.

Then, I conducted a qualitative inductive analysis. First, I identified themes and patterns from the data extracted from the studies, then I summarized the findings and compared them within the studies. I gathered relevant information, organized the data according to the coding sheet and used the AIED taxonomy by Holmes and Tuomi (2022) and (UNESCO, 2021) as guides to compare and identify the AI tools used to support learning in the ESL/EFL K-12 classroom. Although this study is qualitative, I also used logical reasoning to analyze the information from the studies (Howe & Eisenhart, 1990). Thus, I developed a narrative that also presented relevant quantitative information for a better understanding of the topic and to give the reader a perspective of the extent of the effectiveness of AI tools used to support learners in improving their English skills. In addition, I investigated how these AI tools are used in specific

contexts, like inside classrooms with teacher guidance versus independent study outside the classroom. Finally, I used the data collected to create tables, charts, and graphics to illustrate visually the AI tools applied, the findings, the topics, and the limitations related to them.

Chapter 3: Findings

General Information

The general information obtained from the studies selected, like year and country, worked as a reliable source of historical and geographical information to present in this study. The context of countries where AI tools have been developed, along with the different regulations the countries have released guided the perception of how AI tools have been used to support ESL/EFL K-12 classrooms. I was worried about the challenge of choosing only three years for this review, but after reading one systematic review conducted by Alshumaimeri and Alshememry (2024) about AI tools in EFL teaching and learning, I believe I have made a viable choice. The authors included studies between 2007 and 2022 from three academic databases (Web of Science, IEEE, and Scopus). Their focus was to understand the applications of AI tools in the EFL context. They found that 31.5% of the studies about AI tools were published in 2021-2022. They did not include the countries where these tools were used.

Compared to theirs, this work reviewed 23 studies, 4 empirical studies about AI tools in ESL/EFL K-12 were published in 2022, 11 studies were published in 2023, and 8 studies were published in 2024, (see Figure 2). In addition, the studies were conducted in the following countries: Taiwan (8), China (6), South Korea (5), Hong Kong (2), Türkiye (1), and Sweden (1), (see Figure 2).

Another factor that surprised me was the quantity of 21/23 studies of Asian countries. Interestingly, the most famous AI tools (i.e., ChatGPT, Gemini) were developed in the USA, but the studies about their implementation in K-12 were stronger in Asia. I wonder if this

development will come soon in other countries, as they are still considering the potential impact it might bring on society. While some countries are highly concerned about the regulations, others have noticed the powerful impact AI will bring and they have started to test it in all contexts.

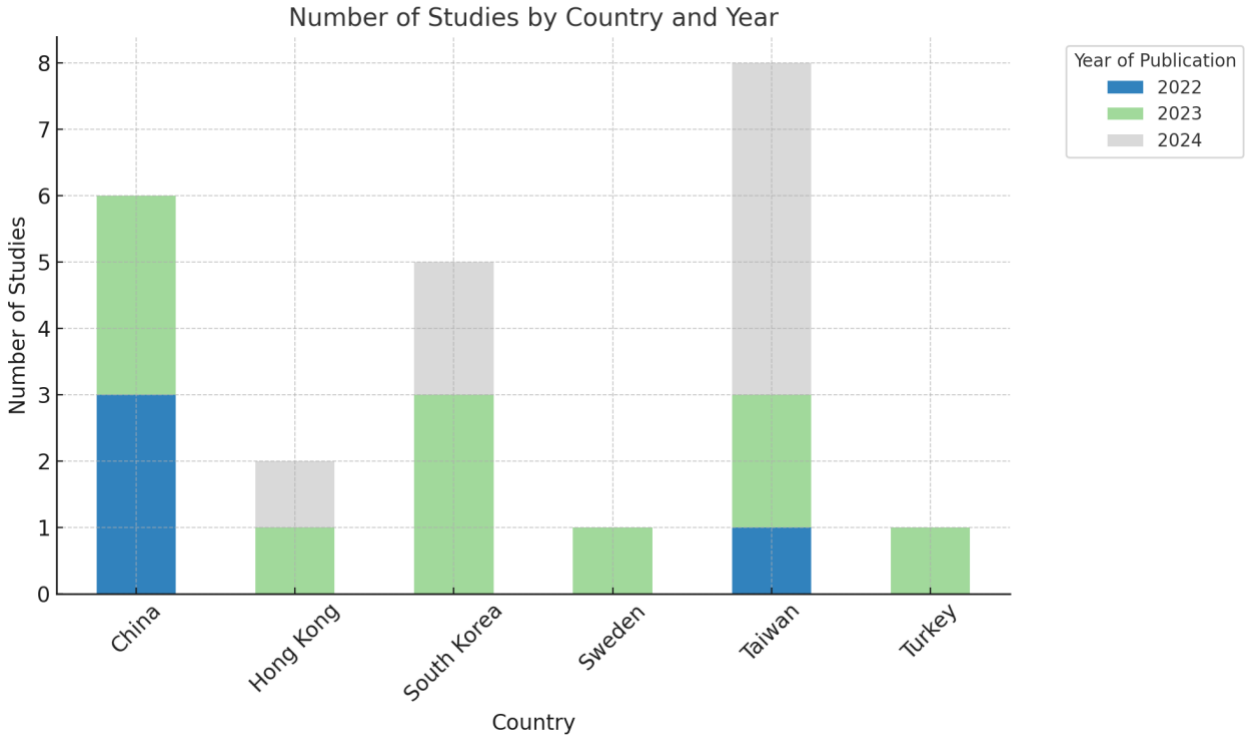
I expected to find studies in the context of ESL because, in countries like the USA and the UK, there are research and discussions about the use of AI tools in ESL K-12 settings (i.e., Diliberti et al. (2024); Holmes and Tuomi, 2022). The RAND report provided information about the districts and schools considering the implementation of AI tools in K-12 education, including language arts, which respond to the needs of most users. However, the RAND report does not present any state or federal recommendations on how schools should include these tools. What happens is that practice comes before research, teachers and students using the tools first.

In all these countries included, English was considered as foreign language (EFL). The average English level of the population in the studies was beginner (A1), elementary (A2), and Intermediate (B1) using the CEFR as a reference. More than 60% of the students were A1-A2 according to the data from the studies. Some studies did not mention the English proficiency level of the students, only the school level. I noticed that the names for school levels changed according to the country, for example, in the studies selected I counted elementary school (6), high school (5), middle school (1), primary school (5), junior high school (2), and secondary school (4). Normally, the use of terms primary school and secondary school are more commonly used worldwide, so to simplify the data collection we might say that in total, it is possible to divide 12 studies in secondary education and 11 in primary education.

Among the studies selected for this analysis, there is one study about perceptions and intentions of teachers (An et al., 2023), one about perspectives of the students about using VH

chatbots (Ericsson et al., 2023), three studies about perceptions of students (Ye et al., 2022; Yang et al., 2022; Tai & Chen, 2023), and 18 studies about AI tools applied in the context of EFL K-12 classrooms. In some experimental and quasi-experimental studies, there were pre-tests, tests, and post-tests. These types of designs are expected from empirical studies. However, some studies presented only a survey after students had contact with AI tools. I also noticed that some authors used the term ESL instead of EFL, but all 23 studies focused on using AI tools in the EFL context in K-12 settings.

Figure 2 *Number of Studies by Country and Year*



Note. This figure is a column chart that compares and shows the number of studies on the vertical axis, the name of the countries on the horizontal axis and the year they were published by color inside the bars with a small text box on the right describing color and year of publication.

Research Question 1: AI Tools Used to Teach EL in ESL/EFL K-12

The AI tools used in the EFL/ESL K-12 context were categorized according to the names mentioned in the studies. It is possible that some authors focused on the functionalities, while others on the categories of AI tools. Before answering research question 1 about the AI tools that were used to teach English learners in K-12 settings, let me offer some information about the

names and types of AI tools. AI tools like chatbots have received more technologies, such as advanced NLP and Machine Learning to improve their functionalities. With these advancements, chatbots like ChatGPT can give more human-like feedback and provide more advanced output for the users when compared to past traditional chatbots like ELIZA which used a basic NLP with specific rules to work and give feedback. Another example is *Duolingo* which is a language learning app that uses advanced AI technology like GPT-4 to improve its conversational applications.

AI tools found in the studies might present some similarities (i.e., personal assistants, chatbots, AI coaches), but they have specific uses in supporting English language learning. For example, the functionalities of a Virtual Human (VH) chatbot might be like ChatGPT, but the VH has an avatar with human characteristics, so the users can interact with this type of tool in a more human-like interaction. Another example is *Praktica* a personal AI avatar tutor created to help people learn English through conversation. The choice of a name might depend on the author's intentions or how the tool is used.

As seen in Table 1 below, I organized 23 studies in alphabetical order and each one received a code from 1 to 23 according to the reference. I categorized AI tools used in the studies by name and type. For the names (e.g., Mondly, Xiaoying, Ellie), I considered the functionalities of the tools, the commercial names they normally have, and the way they were described in the study by the authors. For the types (e.g., chatbot, AI personal assistant), I chose the AI technology involved. In studies where the authors did not mention the names of the AI tools, I chose a name that could identify the type of AI application, for example, Jeon (2024) used a natural language understanding platform called Google's Dialogflow to design a chatbot that would be used in the

EFL context. So, the name I chose was “Chatbot developed with Google’s Dialogflow” and the type was “chatbot”.

Similarly, An et al. (2023) conducted a survey to investigate how AI tools support teaching and learning from the perspective of teachers, referring to the AI tools surveyed in the study as “AI teaching systems”. When I noticed the variety of AI tools in the study, I decided to follow the same categorization “AI teaching systems” for both the name and the type of AI tools. For this process of naming the AI tools, I also considered the fact that there are distinct categories and applications for AI tools used in education. For instance, Ericsson et al., (2023) used Enskill to explore the perspectives of 25 Swedish secondary school students about the use of Virtual Humans (VH). I chose the name to be “Alelo Enskill*”,* and the type “AI chatbot Virtual Human (VH)”, which is an AI-powered embodied chatbot developed to conduct a conversation with the users in a virtual learning environment.

Table 1 *AI Tools Used for ESL/EFL K-12 Learning (RQ1)*

Study Code	Reference	Name of AI tools	Types of AI tools
1	(An et al., 2023)	AI teaching systems	AI Teaching Systems
2	(Chen et al., 2022)	Dynamic Assessment-based Speech Recognition (DA-SR) and Corrective feedback-based speech recognition (CF-SR)	Automatic Speech Recognition (ASR)
3	(Ericsson et al., 2023)	Alelo Enskill	AI Chatbot Virtual Human (VH)
4	(Hsiao & Chang, 2024)	AI-powered tools CooC-Cloud (Linggle Write, Linggle Read, and Linggle Search)	AI Smart teaching assistance platform
5	(Hsu et al., 2024)	AI-IR APP - AI-supported image recognition	AI Image recognition
6	(Jeon, 2023)	Chatbot developed with Google's Dialogflow	AI Chatbot
7	(Jeon, 2024)	Chatbot developed with Google's Dialogflow	AI Chatbot
8	(Kazu & Kuvvetli, 2023)	"Games to Learn English" Artificial intelligence-supported speech recognition WEB 2.0 learning platform assisted by Google	AI Speech Recognition
9	(Lee & Maeng, 2023)	" AI chatbots, including ChatGPT use in English learning"	AI Chatbot
10	(Lee et al., 2023)	AI web-based English learning support system	AI Web-based tool (natural language processing)
11	(Liu & Chen, 2023)	AI-based object detection translation (AI-based ODT)	AI-Based object detection translation
12	(Liu et al., 2024)	Tutee chatbots (Teachable Q&A agent)	AI Chatbot
13	(Rakhun Kim, 2024)	Chatfuel, a code-free chatbot building platform	AI Chatbot
14	(Tai & Chen, 2023)	Google assistant language learning (GALL)	AI Personal assistant
15	(Tai & Chen, 2024a)	Intelligent personal Assistants (Google Home Hub with multimodal feedback and one using Google Home Mini with audio feedback only)	AI Personal assistant
16	(Tai & Chen, 2024b)	Generative artificial intelligence (GAI) Chatbots (CoolE Bot)	AI Chatbot
17	(Wang et al., 2022)	AI coach (Humanized agent)	AI Personal assistant
18	(Wang et al., 2023)	AI coach (Humanized agent)	AI Personal assistant
19	(Woo et al., 2023)	Natural language generation (NLG) tools	Natural Language Processing
20	(Woo et al., 2024)	Prompt engineering (NLG) tools	Prompt engineering
21	(Yang et al., 2022)	Task-based voice chatbot called "Ellie" developed with API platform Dialogflow	AI Chatbot (task-based voice)
22	(Ye et al., 2022)	Chatbot (Microsoft Xiaoying)	AI Chatbot
23	(Yuan, 2023)	Mondly AI chatbot	AI Chatbot

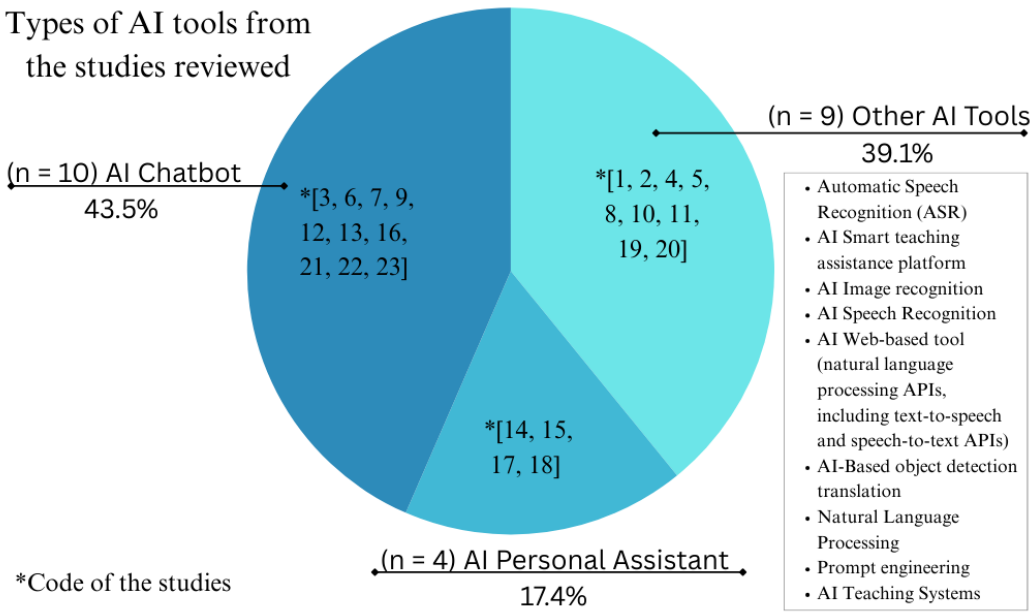
Note. This table describes the AI tools used in ESL/EFL K-12 classrooms. It consists of four columns categorized by study code, reference, name of the AI tools, and types of the AI tools.

In this research, I found various types of AI tools used to teach English in K-12 contexts. Broadly, I identified three types of AI tools used in the studies reviewed: (1) AI Chatbot, (2) AI Personal Assistant, and (3) other AI tools. They were categorized into just three main types to simplify the wide range of AI tools and facilitate the interpretation of the data. First, Chatbots are normally text-based tools that simulate human-like conversation, they provide learners the opportunity to practice the language, ask and answer questions, and simulate dialogs. They account for 43.5% of the tools. AI personal assistants are more voice-driven, they provide multimodal tasks and voice-based engagement, they are usually integrated into a smart device like Google Assistant or Alexa, these tools represent 17.4%, together with chatbots comprising nearly 60% of all AI tools identified. The third category, *Other AI tools* 39.1%, includes those that do not involve conversational interactions. Instead, these tools are primarily designed to provide feedback, assessment, or support input. Unlike chatbots and personal assistants, which emphasize open-ended, human-like interaction, these tools are focused on specific tasks. For instance, when users engage with chatbots, they participate in a simulated dialog, creating a more communicative learning experience. In contrast, tools like AI-based image recognition or AI-based object detection systems do not offer conversational resulting in a different type of learning interaction, it becomes a more targeted and task-specific.

As noted earlier. AI chatbot refers to a type of software used to simulate human conversation, such as *ELIZA* and *ChatGPT.* Whereas AI Personal Assistant refers to a type of software used to give support on daily tasks, such as *Siri,* and *Google Assistant*. The data showed that most recurrent AI tools from the empirical studies are chatbots (see Figure 3). There are 10

studies (43.5% of the studies reviewed) that used chatbots (Ericsson et al., 2023; Jeon, 2023; Jeon, 2024; Lee & Maeng, 2023; Liu et al., 2024; Rakhun Kim, 2024; Tai & Chen, 2024b; Yang et al., 2022; Ye et al., 2022; Yuan, 2023). It is important to highlight that all chatbots in this work are AI-generated and are available on the web or as applications in different app stores. These numbers reflect the recent advancement of AI tools like *ChatGPT, Copilot, Gemini*, and *Deepseek which* is a new AI-generated chatbot launched in January 2025.

Figure 3 *Types of AI tools from the studies*



Note. This pizza chart shows the quantity and types of AI tools divided into three distinct parts by colors (n = 10 AI chatbot, n = 4 AI personal assistant, n = 9 other AI tools). Inside the chart there are the codes of the studies and on the right side a textbox with the names of the AI tools categorized as others.

Four studies (17.4% of the entire studies reviewed) used AI personal assistants. In the studies reviewed, the AI personal assistants (i.e., Google assistant, AI coach) were used to support English language learning. They worked as coaches or personal assistants with the ability to communicate with students, provide interactive practice, and give immediate feedback. Other AI personal assistants like *Siri* and *Alexa* are also extremely popular because of their accessibility. They were not developed to teach English per se, but they have also been used by students and teachers to support their learning (Darda et al., 2024; Tai & Chen, 2023)).

In addition to AI Chatbot and AI Personal Assistant, nine studies (39.1%) used other types of AI tools. For instance, "AI teaching systems" in An et al. (2023) included various AI tools (e.g., chatbots, machine translation), and the authors investigated the perceptions of teachers who used these tools with secondary school students. "Automatic Speech Recognition (ASR)" in Chen et al. (2022) was a system developed by integrating dynamic assessment and speech recognition to enhance the English-speaking skills of elementary school students. "AI Smart teaching assistance platform" in Hsiao and Chang (2024) was a platform called CooC-Cloud, which was empowered by three AI tools (Linggle Write, Linggle Read, and Linggle Search) launched by the city of Taipei City Department of Education to support students to read and write. "AI Image recognition" in Hsu et al. (2024) was an AI-supported image recognition technology used to support students in learning vocabulary in a multi-sensory experience.

Similarly, "AI Speech Recognition" in (Kazu & Kuvvetli, 2023) was an Artificial intelligence-supported speech recognition platform used to teach vocabulary through pronunciation to high school students. "AI Web-based tool (natural language processing)" in Lee

et al. (2023) was an AI-applied system developed with Google Cloud Natural Language API, Twinword's Word Dictionary API, Amazon Polly, and other AI technologies guided by the Learner Generated Context (LGC) framework to assist learners in studying English autonomously. "AI-based object detection translation" in Liu and Chen (2023) was an application developed to facilitate vocabulary learning of elementary school students through the presentation of objects in picture, word, and pronunciation formats. "Natural Language Processing" in Woo et al. (2023) was related to four Natural Language Generation (NLG) tools developed using *Hugging Face* to support secondary school girls in creative writing by facilitating their idea generation skills. Finally, "Prompt engineering" in Woo et al. (2024) was used to help secondary school students to develop NLG tools and use them to support their story-writing activity.

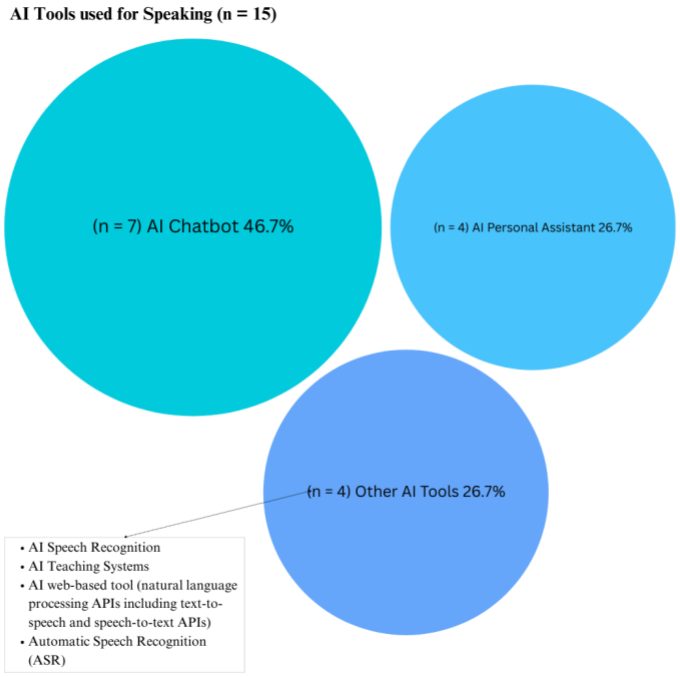
Research Question 2: What Language Skills Were Taught with AI Tools

AI tools are dynamic and can be used to develop multiple language skills at the same time. However, the analysis of this study was informed by the researchers' reports on the language skills they examined using AI tools. For instance, if the authors reported that they investigated a specific AI tool to develop skills, such as 'speaking' and 'listening', these two skills would be the ones marked for the analysis, even though the AI tools could also be used to develop reading and writing skills.

Among the four English skills supported by AI tools in the 23 studies reviewed, 'Speaking' skills are most frequently explored. Speaking, reading, writing, and listening skills were examined in 15, 10, 8, and 7 studies, respectively. This finding is quite unexpected because in higher education reading and writing skills were the most examined and/or best supported with AI tools, but unexpectedly, in K-12 speaking skills were apparently most studied. Among the 15

studies on speaking skills, seven studies used AI chatbots (46.7%), four studies (26.7%) used AI personal assistants, and another four studies used other AI tools (i.e., AI Speech Recognition, AI Teaching Systems, AI web-based tool, and Automatic Speech Recognition) to examine the development of speaking skills, (see Figure 4).

Figure 4 *Types of AI tools used for supporting speaking skills*



Note. This figure shows three proportional circles in assorted colors with the percentage of types of AI tools used for supporting speaking skills.

The most recurrent aspects of speaking skills covered by the studies include pronunciation, vocabulary, speaking, conversation, practice, proficiency, willingness to communicate, shadowing, and oral practice, (see Table 2). These findings about specific aspects

of speaking skills describing how teachers and students can use AI tools to support speaking are important because they can be used to support the practice inside and outside the classroom. In some studies (e.g., Chen et al., 2022; Ericsson et al., 2023; Kazu & Kuvvetli, 2023; Tai & Chen, 2024a, 2024b; Ye et al., 2022; Yuan, 2023), the authors conducted pre-tests, tests, and post-tests to measure the efficacy of these tools to support speaking and they showed large effect size. In addition, Yuan (2023) and (Tai & Chen, 2024b) stated that chatbots could facilitate and mediate language learning instruction.

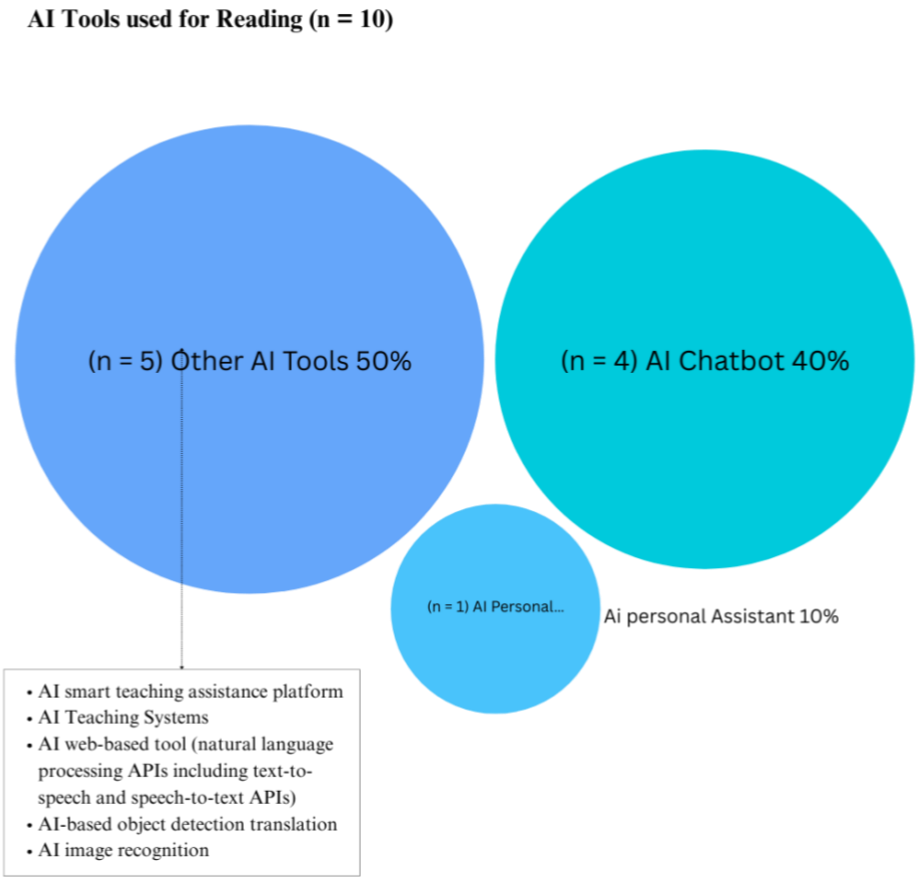
Table 2 *AI Tools Used for Speaking Skills*

Study Code	Type of AI tools	Speaking Skills
1	AI Teaching Systems	Perceptions
2	Automatic Speech Recognition (ASR)	Facilitate speaking - Tasks: Picture reading, sentence pattern reading, and short conversations.
3	AI Chatbot Virtual Human (VH)	Interactional skills, satisfaction, Engagement
7	AI Chatbot	Affordances
8	AI Speech Recognition	Pronunciation (vocabulary acquisition) word retention,
9	AI Chatbot	Perceptions
10	AI web-based tool APIs)	Practice
14	AI personal assistant	Willingness to communicate, pronunciation, vocabulary
15	AI personal assistant	Proficiency (reading a conversation, answering questions, describing a picture)
16	AI Chatbot	Speaking practice (Individual and paired interactions)
17	AI personal assistant	Pronunciation problems, develop vocabulary, and conversational skills, English shadowing
18	AI personal assistant	Shadowing, pronunciation, vocabulary
21	AI Chatbot (task-based voice)	Vocabulary, conversation
22	AI Chatbot	Oral tasks (English accuracy, grammar, and pronunciation)
23	AI Chatbot	Oral English proficiency, Willingness to communicate

Note. This table brings three columns with the code of the study, the types of AI tools, and the speaking skills.

Reading is another skill that was frequently researched and supported by AI tools. Among 23 studies, ten studies examined reading skills development, and AI tools. Between these ten studies, four studies (Jeon, 2023, 2024; Lee & Maeng, 2023; Liu et al., 2024) used AI Chatbot, another five studies (An et al., 2023; Hsiao & Chang, 2024; Hsu et al., 2024; Lee et al., 2023; Liu & Chen, 2023) used other AI tools, such as AI teaching systems, AI-IR APP - AI-supported image recognition, AI smart teaching assistance platform, AI image recognition, AI web-based tool (natural language processing APIs including text-to-speech and speech-to-text APIs), AI-based object detection translation and only one study (Wang et al., 2023) used AI Personal Assistant as seen in Figure 5.

Figure 5 *Types of AI tools used for supporting reading skills*



Note. This figure shows three proportional circles in assorted colors with the percentage of types of AI tools used for supporting reading skills.

The data show that various AI tools were used to help ELs develop their reading skills. As seen in Table 3, these ten studies focused on developing reading-related skills. In this study, reading is broadly conceptualized, thus reading skills in this study include reading practices, reading sentences, reading interests, reading engagement, vocabulary performance, vocabulary acquisition, as seen in the right column in Table 3.

Table 3 *AI tools used for reading skills*

Study Code	Type of AI Tools	Reading Skills
1	AI Teaching Systems	Perceptions
4	AI smart teaching assistance platform	"Grammatical and word- choice. Reference information of collocations and grammar patterns on demand."
5	AI image recognition	Vocabulary performance and acquisition, anxiety and behavior
6	AI Chatbot	Vocabulary learning (glossing) receptive and productive
7	AI Chatbot	Affordances: Immediate feedback, reduce anxiety, self-regulation
9	AI Chatbot	Perceptions: Benefits, concerns, ethical considerations
10	AI web-based tools (natural language processing APIs including text-to-speech and speech-to-text APIs)	Practice (texts, vocabulary)
11	AI-based object detection translation	Vocabulary (image, text, pronunciation)
12	AI Chatbot	Reading Interest and Engagement
18	AI personal assistant	Reading sentences

Note. This table brings three columns with the code of the study, the types of AI tools and the reading skills.

One important finding is that vocabulary learning with AI tools was quite evident in studies on reading development (i.e., Hsu et al., 2024; Jeon, 2023; Lee et al., 2023; Liu & Chen, 2023). The use of AI tools for vocabulary acquisition might be a good takeaway from the studies reviewed. L2 studies have shown the critical importance of vocabulary learning for reading development (Krashen, 1989; Nation, 2001; Ortega 2009). In addition, according to the National Reading Panel (2000) vocabulary is one of the five essential components for reading skills development, the others are phonemic awareness, phonics, fluency, and comprehension. Thus, AI

tools can easily support students on tasks like choosing words and asking the AI tools to generate texts and other reading-related activities. These capabilities and applications of AI tools show the ways they may support students in interacting and engaging with texts and vocabulary as a key element in the learning process of English learning.

The analysis of the ten studies on the use and perceptions of AI tools for reading revealed that chatbots were the most commonly used tool to enhance reading skills, appearing in four studies (i.e., Jeon, 2023, 2024; Lee & Maeng, 2023; Liu et al., 2024) (see Figure 5 and Table 3). The popularity of Chatbots like ChatGPT might be one of the reasons for this. Lee and Maeng (2023) stated that 53.3% of the participants of their study had already used ChatGPT because it can work with multilingual communication, content creation, translation, and other useful applications that were not available in previous chatbots. Jeon (2024) showed that chatbots have affordances, such as providing immediate feedback, reducing anxiety, and supporting self-regulation. In addition, students can stop the chatbot at any time during the practice to search for more information about unknown words or topics.

Specifically, about reading, the studies focused on the interactive capabilities and vocabulary acquisition for reading improvement. Liu et al. (2024) focused on engaging students to help them to improve their reading skills. To do so, they implemented the learning by teaching pedagogy. Their experiment consisted of 95 fifth-grade students divided into two groups, 47 students were randomly assigned to an experimental group and the other 48 remained as the control group. The experimental group received a tablet to engage with chatbots in battling and training. Students were supposed to ask the chatbot questions about the book they read and invite

their peers to test which chatbot had learned more about after the training. This experiment enhanced the students' reading interest and engagement.

Jeon (2023) integrated assessment into vocabulary instruction, he focused on receptive and productive vocabulary and used a chatbot that was created with Google Dialogflow. He had two experimental groups of primary school students; one group received graduated assistance from a chatbot, and the other group received target word definitions. The control group did not receive any support from the chatbot. The three groups were asked to read texts and then identify the meaning of underlined target words. He assessed the students through pretest, posttest, and delayed posttest. The results showed that students from the two experimental groups using chatbots had higher gains in vocabulary and reading than the students in the control group. The author also conducted a qualitative analysis and found that students had developed their vocabulary within ZPD (Zone of Proximal Development), which might be a good indication of progressive learning for the students.

In addition to AI chatbots, other AI tools were used to support reading development. More specifically, other AI tools such as AI Teaching Systems, AI smart teaching assistance platform, and AI web-based tools, were used to support students in word choice, grammar, practice, pronunciation, and collocations. It is also interesting to see technologies like image recognition and object detection as useful resources for developing reading skills. It is possible to see the following aspects, such as vocabulary, grammar, and practice (see Table 3). Vocabulary was again an important aspect covered by 3 studies and it is connected to An et al. (2023) who found

that the perceptions of teachers about using AI teaching systems to help students learn vocabulary and grammar were positive.

Besides the perceptions of teachers, Hsu et al. (2024) used AI image recognition to support primary school students in vocabulary learning. In this study, the focus is on diverse dimensions of vocabulary learning including its connection to reading skills. The authors highlighted that words and phrases students learn in textbooks should be connected to their reality, and that is when AI tools like Image recognition are important because they can create a kind of interdependence between textbook-introduced vocabulary and real-world cognition. For the learning task, they chose "Tom's Daily Routine" a text passage commonly used in EFL/ESL textbooks.

The students needed to go around the classroom with a mobile powered with an AI image recognition app to scan the images and learn the words connected to the text. Liu and Chen (2023) examined how AI-based ODT (object detection translation) could affect students' vocabulary learning and their experiment was like that of Hsu et al. (2024). They started with a warm-up story-reading exercise before the experiment. Then they asked students of the experimental group to use their smartphones powered with the AI-based ODT and scan objects such as fruits and vegetables, the app presented them in text, picture, and pronunciation formats. The control group used Google Translate by typing the name of the object in Chinese. After the experiment, the posttest showed that students using AI object detection had a better vocabulary improvement when compared to the control group.

Interestingly, when compared to other skills (e.g., 4 studies for speaking skills) AI personal assistant was used in only one study for reading instruction. It is difficult to determine

why, but only one study (i.e., Wang et al., 2023) examined the use of AI Personal Assistant for English shadowing through sentence reading. In this study, the AI personal assistant reads a sentence from a textbook, and the student repeats the sentences, if the AI personal assistant identifies any problem, it scores the students and gives feedback by using supportive statements like 'I am so proud of you' or 'this sentence is incorrect". Normally, shadowing is used for speaking and listening, however, it is observed that students read actively during the experiment.

Next to speaking and reading, writing skills were examined and supported with the use of AI tools in eight studies (An et al., 2023; Hsiao & Chang, 2024; Jeon, 2024; Lee & Maeng, 2023; Lee et al., 2023; Rakhun Kim, 2024; Woo et al., 2023; Woo et al., 2024). A wide range of writing skills were examined and supported in these studies. For instance, An et al. (2023) showed that the perception of teachers about the use of AI teaching systems to analyze the meaning of sentences, and detecting grammatical structure was positive. That view is in line with Hsiao and Chang (2024) who examined AI smart teaching assistance platform with the support of three AI-powered tools (i.e. Linggle Write, Linggle Read, and Linggle Search) to support writing and cultivate the aspect of autonomy of the EFL learners. They focused on Grammatical Error Correction (GEC), identifying error-prone content words and collocations, and providing corrective feedback.

Additionally, Jeon (2024) stated that receiving corrective feedback, which is helpful to help ELs identify and correct their mistakes while they are writing, is one of the advantages of using AI chatbots. When drafting stories, people need to write, review, and publish, but this process can involve other steps and skills, such as a continuous process of writing, correcting, and rewriting. These aspects of writing practice, corrective feedback, and recast were prevalent when

students were writing their pieces (Rakhun Kim et al., 2024). The AI tools analyzed in these studies gave the students continuous feedback and suggestions on how to improve their writing. It was like what students normally receive from other AI tools. Jeon (2024) demonstrates how students engage and learn deeply when they develop their own AI tools to learn languages. It looks like the students feel part of the creative process, and then they test the AI tool and learn at the same time.

One finding from the tools used for writing was the focus on grammar, specifically corrective feedback, and recast (see Table 4). In language acquisition, several authors have discussed corrective feedback (CF). For instance, Krashen (1985) argues that CF is not necessary, as it may cause anxiety, though recasts can be useful if they do not interfere with communication for ELs. Swain (1985, 2005) suggests that CF helps students refine their output, while recast alone is insufficient if it does not encourage the learner to change the output. Long (1996) views CF as beneficial when it facilitates negotiation meaning. He emphasizes that active learner engagement should be accompanied by recast to enhance its effectiveness.

Finally, Lightbown & Spada (1990) discuss CF on the connection between communicative language teaching and grammar learning. When at school, students need to think about a topic, or they are given one by their teachers according to the content, then they brainstorm some ideas and start to write them. Lee et al. (2023) developed an AI web-based tool that incorporated natural language processing (NLP) APIs, including text-to-speech and speech-to-text functionalities to support ELs in multi-sensory activities, such as speaking, listening, and writing to support students in conducting all the process autonomously and connected to their specific needs. In

writing during this activity students could practice by visualizing and typing sentences and words in the text mode of the tool. The focus was on sentence and word writing.

Table 4 *AI tools used for writing skills*

Study Code	Type of AI Tools	Writing Skills
1	AI Teaching Systems	Perceptions
4	AI smart teaching assistance platform	Grammatical Error Correction (GEC), error-prone content words and collocations, corrective feedback
7	AI Chatbot	Affordances (corrective feedback)
9	AI Chatbot	Perceptions
10	AI web-based tool (natural language processing APIs including text-to-speech and speech-to-text APIs)	Practice
13	AI Chatbot	Automatic corrective recast (learning of the English caused-motion construction) Grammar
19	Natural Language Processing	Write an English language short story
20	Prompt engineering	Write a story no more than 500 words

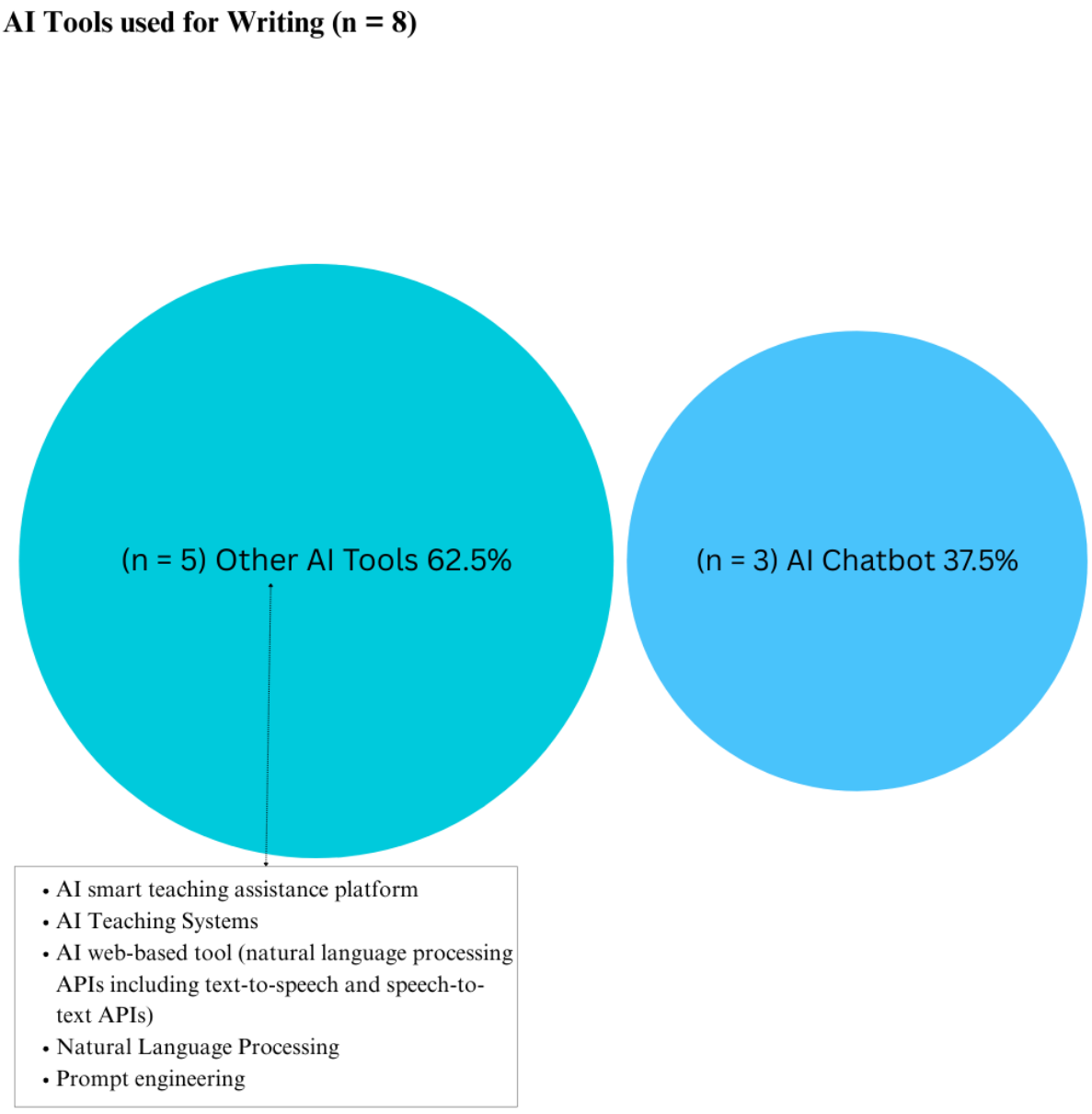
Note. This table brings three columns with the code of the study, the types of AI tools, and the writing skills.

Among the eight studies on writing development with AI tools, only 3 studies (i.e., Jeon, 2024; Lee & Maeng, 2023; Rakhun Kim, 2024) used chatbots; the rest of five studies (i.e., (An et al., 2023; Hsiao & Chang, 2024; Lee et al., 2023; Woo et al., 2023, 2024) used other AI tools (see Figure 6). Among the other AI tools, five distinct types can be identified, including Natural

Language Processing (NLP) and Prompt Engineering (PE). Notably, NLP and Prompt engineering were connected to Natural Language Generation (NLG) in two studies (i.e., Woo et al., 2023, 2024). Normally, NLP is what makes communication between humans and machines possible, it makes the AI model process and respond to human language. Prompt Engineering (PE) is how to find a better way to communicate with the machine, and NLG is a subfield of NLP that generates a more human language connected to a context. NLP and PE are technologies that could be included in AI applications such as chatbots.

However, in these studies, they were used to support the students in developing and analyzing their own AI tools while writing in English. Woo et al., (2023) developed four NGL tools on Hugging Face and focused on idea generation for high school students writing. Students needed to write a text, submit it, and receive ideas on how to improve their writing. Woo et al. (2024) focused more on creating NLG tools and using them to draft a story. In both studies, the students needed to draft a story of 500 words or more and the authors suggested the use of these tools for improving the quality of their stories.

Figure 6 *Types of AI tools used for supporting writing skills*

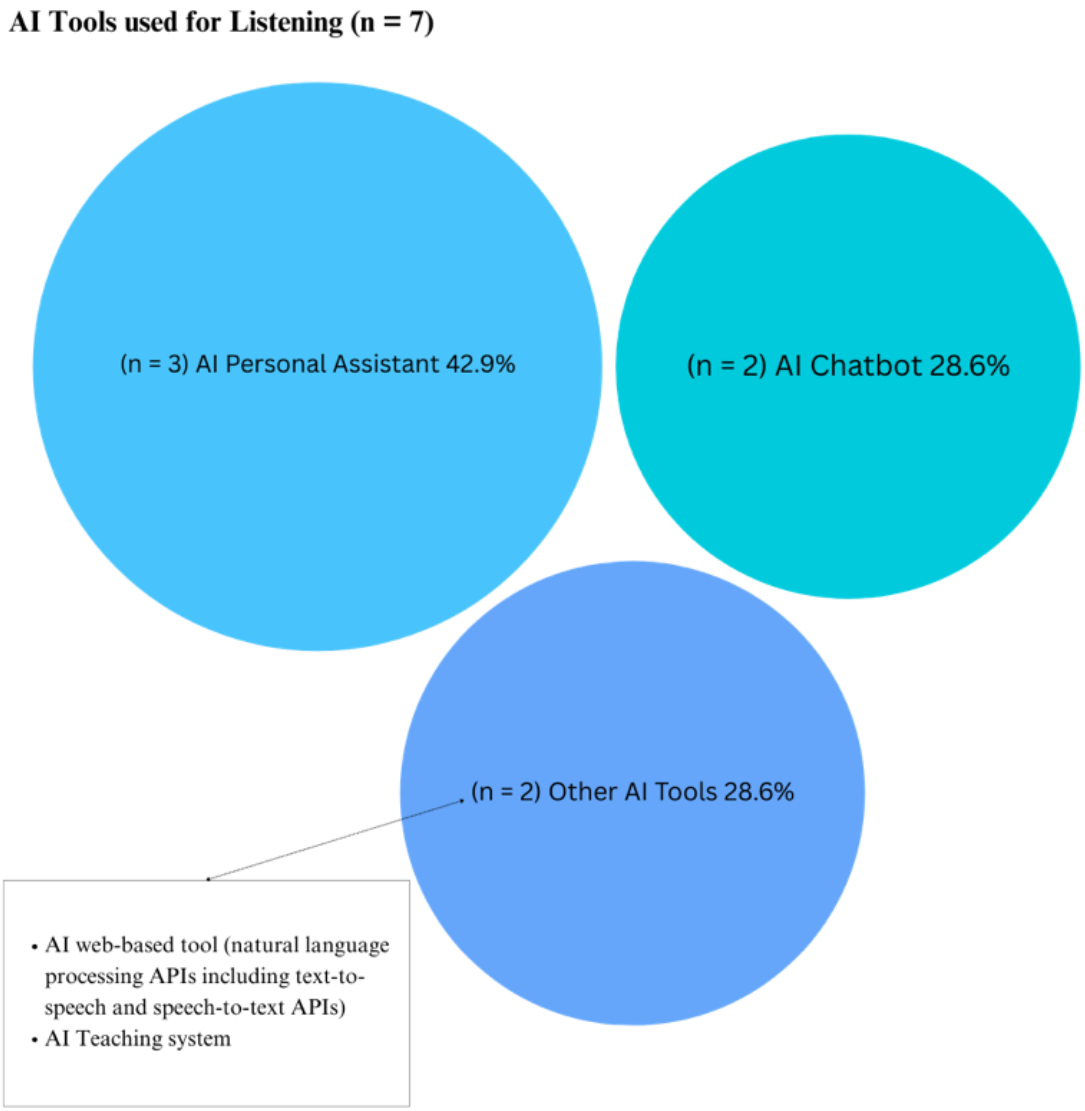


Note. This figure shows two proportional circles in assorted colors with the percentage of types of AI tools used for supporting writing skills.

Finally, listening skills were least examined and supported by the AI tools. Among 23 studies reviewed, only seven studies (i.e., An et al., 2023; Jeon, 2024; Lee & Maeng, 2023; Lee et al., 2023; Tai & Chen, 2023; Wang et al., 2022; Wang et al., 2023) focused on listening skills with AI tools. One interesting finding is that chatbots were not the most popular, but AI personal assistants (see Figure 7). This finding is connected to the history and advancements of chatbots, which primarily were developed for conversational interaction through text. This name chatbot came from ChatterBot and was coined by Mauldin (1994). An et al. (2023) mentioned that the perceptions of teachers about using AI teaching systems, including chatbots empowered with speech recognition technology for pronunciation accuracy, fluency, and phonological tone were positive for the EFL classroom. About the perceptions of students Lee and Maeng (2023) highlighted that Chatbots were perceived by the students as easier to understand and to use.

Using Google Dialogflow, Jeon (2024) created chatbots to support students and identified pedagogical affordances such as increased listening opportunities and immediate feedback. Students responded positively to their experience with extended and goal-oriented conversations. While these chatbots provided speech-based interaction the study primarily focused on chatbot affordances. It is important to notice that modern chatbots have advanced significantly and they can match the technology present in AI personal assistants that are specifically developed for interactive listening and speaking, but it will depend on the design and purpose of the AI tool.

Figure 7 *Types of AI tools used for supporting listening skills*



Note. This figure shows three proportional circles in assorted colors with the percentage of types of AI tools used for supporting listening skills.

AI personal assistants were not used for supporting writing skills in the studies reviewed but were most frequently used for listening skill development. As I mentioned before, AI personal assistants are primarily developed for listening and speaking. According to Tai and Chen (2023) and Wang et al. (2022, 2023) using personal assistants for listening comprehension has worked well, and it is the aspect most covered in the studies (see Table 5). The use of AI personal assistants in Wang et al. (2022, 2023) involved speech recognition and NLP to assess the students' output and give feedback during the listening practice. The feedback was given through AI's linguistic models. In (Tai & Chen, 2023) AI personal assistant was used as a coach, the students had two tasks to practice their listening, they first were supposed to enhance their listening comprehension by listening to a story, a poem, or facts related to their preferences and the second was listening to real music and engage in a trivia game.

Table 5 *AI tools used for listening skills*

Study Code	Type of AI Tools	Listening Skills
1	AI Teaching Systems	Perceptions
7	AI Chatbot	Affordances
9	AI Chatbot	Perceptions
10	AI web-based tool (natural language processing APIs including text-to-speech and speech-to-text APIs)	Practice
14	AI personal assistant	Listening comprehension
17	AI personal assistant	Listening comprehension
18	AI personal assistant	Listening practice

Note. This table shows which AI tools are used for specific listening skills.

The use of AI tools to support the four English skills (reading, writing, listening, speaking) was discussed in the 23 studies selected and it was possible to see a prevalence of their use for reading and speaking skills development. However, it is important to point out that one of the aspects presented in most experiments was vocabulary. It looks like vocabulary plays a central role in connecting different domains and several AI-driven tools. This is completely aligned with this study because it focuses on primary and secondary school students learning English as a second language, and as seen in the 23 students the students were at a basic level of proficiency in English, which makes sense the focus on learning vocabulary and sentences.

Research Question 3: Perceptions of Teachers and Students about AI Tools

In this study, I examined teachers' and students' perceptions of using AI tools to teach and learn English as an additional language. Three major aspects of their perceptions were identified. Thus, I report their perceived (1) effectiveness, (2) challenges, and (3) limitations in the order.

Perceived Effectiveness

The perceptions of students and teachers about the effectiveness of AI tools to support English language learning in EFL/ESL K-12 classrooms were positive in 11 studies included in this work. Among the 11 studies, just three studies (i.e., An et al., 2023; Ericsson et al., 2023; Yuan, 2023) focused on reporting teachers' perceptions. After conducting pretests and posttests, Yuan (2023) carried out a semi-structured interview with ten questions for two teachers and seven questions for focus groups with 74 students, both groups agreed chatbots improve oral English proficiency.

An and her colleagues examined the perceptions of 470 middle school EFL teachers in China concerning the effectiveness of AI tools in their EFL K-12 classrooms. In this survey, the participating teachers were asked to answer a questionnaire divided into two parts, the first was to collect their demographic information and the second was to measure eight key concepts (1) Performance Expectancy (PE), (2) Effort Expectancy (EE), (3) Social Influence (SI), (4) Facilitating Conditions (FC), (5) AI language technological knowledge (AIL-TK), (6) AI technological pedagogical knowledge (AI-TPK), (7) AI technological pedagogical content knowledge AI-TPACK, and (8) Behavioral Intention (BI). Among these concepts, (1) Performance Expectancy was the most relevant topic used to measure the effectiveness of AI tools. The data showed that teachers had high Performance Expectancy toward the use of AI tools in their teaching practice when compared to the other seven concepts, and perceived AI teaching systems as effective for teaching quality, efficiency, and overall usefulness.

An et al. (2023) defined Performance Expectancy as an individual's belief in how much a system can enhance their work performance. In this study, all participating teachers had prior experience with AI tools. They were asked to indicate their level of agreement on a 5-point Likert scale (1 = Strongly Disagree to 5 = Strongly Agree) with the following statements: (1) AI can help me improve the quality of teaching, (2) AI can help me improve the efficiency of teaching, (3) I believe AI is very useful in my job, (4) AI is very helpful for my teaching. The positive 4.3 average score indicated that teachers who answered the survey were not concerned about the possibility of implementing AI tools to support their L2 instructions. The study demonstrated that this perception of AI tools being useful for improving their teaching performance is important because it influences their Behavior Intention (BI) to use technology in the classroom context. In

summary of the other seven concepts the means scored above three and showed that teachers have from neutral to positive perceptions, knowledge about AI tools, and behavioral intention to use AI tools in EFL settings.

In addition, seven studies (i.e., Ericsson et al., 2023; Jeon, 2024; Lee & Maeng, 2023; Tai & Chen, 2023; Wang et al., 2023; Yang et al., 2022; Ye et al., 2022) examined the perceptions of the students about the effectiveness of using AI tools to learning English as an additional language.

Four studies (i.e., Jeon, 2024; Lee & Maeng, 2023; Tai & Chen, 2023; Wang et al., 2023) examined the students' perceptions about improving the four English skills by using AI tools, whereas three studies (i.e. Ericsson et al., 2023; Yang et al., 2022; Ye et al., 2022) focused on reporting students' perceived effectiveness of using AI tools for improving one specific language skills. For instance, Yang et al., (2022) used a task-based voice chatbot called "Ellie" to assess the oral skills of students. *Ellie* worked as a partner to improve the English-speaking skills of the students. The students highlighted how they felt encouraged by engaging in conversation with the chatbot. The idea of measuring the efficacy of AI tools through the students' engagement has been a positive aspect indicator for a smooth implementation in the classroom, according to Yang et al., (2022) it is not common to observe this engagement of Korean students in conversation. The results showed that the students evaluated the use of *Ellie* chatbot as positive.

Similarly, Ye et al. (2022) analyzed the perceptions of the students in China about using Microsoft Xiaoying chatbot to learn English as a foreign language. The students were asked about their perceptions of using this tool on smartphones to improve oral accuracy. The results showed they perceived improvement in their oral performance, grammar, and pronunciation accuracy

using the chatbot. What's notable is that although Yang et al (2022) and Ye et al (2022) studies implemented two different chatbots (Ellie in Yang et al's study in 2022 and Xiaoying in Ye et al, 2022), in the same year, in different countries, they both revealed students' positive perceptions of using AI tools to develop speaking skills, including grammar, conversation, pronunciation, engagement, and motivation.

Jeon (2024) addressed one relevant finding about students' perceptions of the social affordances of using chatbots in EFL classrooms for learning English. Among the participants, 26 students appreciated interacting with chatbots as it allowed them to control their learning pace without peer pressure. In contrast, 10 students preferred studying with peers instead of chatbots. These findings suggest that students' views on the effectiveness of AI tools are influenced by their social interaction preferences. In other words, social preferences play a role in their perceptions of chatbot effectiveness.

Jeon (2024) also identified pedagogical affordances (i.e., immediate feedback) and technological affordances (i.e., a stop button to control the pace) in the use of chatbots. His exploratory study involved 36 Korean primary school students with novice-level English proficiency for 16 weeks (about 3 and a half months). During the first two weeks, students were introduced to the chatbots, while the remaining 14 weeks (about 3 months) consisted of instruction delivered through whole-group, small-group, and individual conversation with chatbots to support tasks, such as target words, sentences, games, songs, and chants. After the tasks, students were interviewed on the following questions (1) Can you describe how you used the chatbots? (2) Did you find using chatbots helpful when studying English? Why? (3) What difficulties did you have when studying with chatbots? (4) How did you overcome these

difficulties? (5) Do you want to continue to use chatbots in English class? Why? The interview showed that in general students had positive perceptions about using chatbots to complete their tasks and get motivated to interact with them in English.

In addition to English learners' social perceptions about interacting with chatbots described by Jeon (2024), the type of *interaction* influences the perception of the students. For instance, Wang et al. (2023) investigated the perceptions of 327 primary school students about the presence of an AI coach that worked as a web-based chatbot. The students used an AI agent, then they gave their perceptions about its use. This chatbot was an avatar acting like a humanized agent. The authors found that students' perceptions of AI tools' presence might influence their enjoyment and outcomes.

Teachers' and students' perceived effectiveness on the use of AI tools in EFL classrooms indicated that, while they recognized the pedagogical and technological support these tools provide, the human aspect of the interaction between teachers, students, and chatbots around the content they are supposed to learn may influence the future directions of AI tools in education.

Perceived Challenges

Despite the perceived effectiveness of using AI tools, teachers expressed challenges of using AI tools to support their instruction (An et al., 2023; Ericsson et al., 2023; Yuan, 2023). The three most frequently perceived challenges are (1) the lack of support, (2) the lack of professional development, and (3) the lack of teacher knowledge.

The first challenge teachers perceived was *the lack of support* they received from stakeholders. To address this challenge An et al., (2023) used a construct called Facilitating Conditions (FC), which measures how much the teachers believe they have an organizational and

technical infrastructure to support their use of AI tools in teaching. In this study teachers' perceptions of Facilitating Conditions (FC) scored 3.7 on a Likert scale (1 = disagree, 5 = strongly agree), suggesting a neutral to slightly positive view of the support. The authors considered this result a challenge because the survey was conducted exclusively with teachers from an AI demonstration area in China, and it requires further verification. This finding aligns with the results of other surveys. For instance, Diliberti et al. (2024) stated that 38% of teachers in the USA who used AI tools in K-12 education reported a lack of guidance from their districts regarding the use of AI tools, which would correspond to approximately 3.5 of FC.

This lack of support leads to the second challenge, *the lack of professional development* about using AI tools to teach English. An et al. (2023) stated that FC has a strong relationship to professional development. They stated that professional development should provide teachers with the necessary guidance and support for AI-supported teaching to design and implement classroom activities where the students can participate actively in their learning process. Yuan (2023) mentioned that teachers may be resistant to implementing AI tools in their teaching practice because they are not familiar with how to do it, or because they do not know about the efficacy of AI tools. In the USA, Diliberti et al. (2024) stated that 35% of the teachers who used AI tools in K-12 education reported a lack of professional development from their districts. These numbers show little difference in the lack of professional development between high-poverty and low-poverty schools.

This second challenge, the lack of professional development, also seems to affect the third challenge: *the lack of teacher knowledge.* Ericsson et al. (2023) reported teachers perceived constraints because the AI tools were not aligned with the Swedish curriculum. These perceptions

emerge when teachers know their curriculum, the content they are supposed to teach, and the technologies they are supposed to use. An and her colleagues divided teacher knowledge into three parts; (1) AI Language Technological Knowledge (AIL-TK), which refers to teachers' knowledge about AI technology and its application for language learning and instruction, (2) AI technological pedagogical knowledge (AI-TPK), which involves the knowledge about how teaching changes due to integration of AI technologies, and (3) AI technological pedagogical content knowledge (AI-TPACK), which represents the knowledge about teaching with AI technology and subject matter.

Two studies (An et al., 2023; Yuan 2023) revealed that teachers perceived the lack of AI technological pedagogical knowledge (AI-TPK) as a challenge. The survey showed an average score lower than AIL-TK, and AI-TPACK). This challenge influences the way teachers design their instruction, manage their classrooms, and engage the students. An et al. (2023) measured these aspects using a 5-point Linkert scale for the following statements: (1) I know how to use AI tools to plan courses, (2) I know when I should use AI in teaching, (3) I know how to use AI tools to set students' learning goals. An et al. (2023) indicated AI teachers' knowledge average score is 3.5, which means teachers were neutral and not as positive as they were when assessing the effectiveness of AI tools in their performance expectancy. This result is slightly positive, but it still indicates that teachers have faced challenges related to the lack of AI knowledge, this affects directly on how they use the strategy of personalized guidance to improve the skills of their English learners with the help of AI. Yuan (2023) asked how teachers could remodel themselves to integrate AI chatbots into classrooms better; the answers showed a skeptical and slightly resistant attitude about it.

In addition to challenges teachers face, students' perceived challenges of using AI tools were examined in eleven studies (i.e., Ericsson et al., 2023; Jeon, 2024; Kazu & Kuvvetli, 2023; Lee & Maeng, 2023; Tai & Chen, 2023, 2024a, 2024b; Wang et al., 2023; Yang et al., 2022; Ye et al., 2022; Yuan, 2023). Interestingly, the perceptions of students about the challenges related to AI tools are different from the teachers' perceptions because the way they interact with these tools seems to be quite different from the way teachers do. The students' perceptions are related to their individual experiences and how they cover their specific needs using AI tools. The two most frequently perceived challenges are (1) the technical limitations of AI tools, and (2) the lack of flexibility.

The first challenge is related to technical issues. For instance, in three studies (Ericsson et al., 2023; Jeon, 2024; Lee & Maeng, 2023) the perceived challenge presented by the students was the fact that the AI tools used did not understand what they had said. This situation made them feel frustrated. In addition, they found that girls had more difficulty being understood by the Chatbot. Kazu and Kuvvetli (2023) conducted a study with 56 high school students and asked the experimental group (n=28) in which ways they had difficulty using AI tools, and the challenges students perceived were related to their memorization of words. That might be related to what Tai and Chen (2023) mentioned about the challenge students have in tracking their progress when they use AI personal assistants. For Kazu and Kuvvetli (2023) and Lee and Maeng (2023), the challenges perceived by students were related to the fact that the chatbot stopped working during the tasks. And other technical issues mentioned by Yang et al., (2022) about problems related to chatbot's voice recognition in other cases the lack of more vocabulary impeding learners' oral performance Ye et al., (2022).

The second challenge is the lack of flexibility. Due to the need for more flexibility for differentiated learning, Chatbots face linguistic and comprehension obstacles including advanced vocabulary, lengthy sentences, and content that exceeds age-appropriate levels (Tai & Chen, 2024b). Additionally, they struggle to accurately recognize accented speech (Tai & Chen, 2024a). Yuan (2023) also discussed the challenge of attracting and maintaining learners' interest. Wang et al., (2023) emphasized how the students' perceptions about the appearance of AI avatars influenced L2 their engagement and how this should be a topic to be covered by the development of future AI tools.

Perceived Limitations of using AI tools

Limitations of AI tools are similar or related to the challenges that teachers and students face as noted above. My analysis reveals three key limitations from the 23 studies reviewed: ethical issues (e.g., academic misconduct), data privacy concerns (e.g., personal data leaks), and over-reliance on AI tools. Among 23 studies reviewed in this study, none of the studies directly examined teachers' perceptions about the limitations of AI tools used in the context of ESL/EFL K-12 classrooms. Only one study (Lee & Maeng, 2023) examined students' perceptions of the limitations of using AI tools. They conducted a survey with thirty high school students from Korea to explore their perceptions of using chatbots for English learning, including benefits, concerns, and ethical issues. They found that such limitations are related to ethical and privacy concerns. Students demonstrated high concerns about ethical issues, these limitations affect the way students use their creativity and originality, and avoid potential infringements of copyright, such as plagiarism. Students were also concerned about data privacy, they believed AI tools

would use their personal information without permission, record their conversations, or misuse their data for marketing.

In addition, Lee and Maeng (2023) found that Korean high school students were concerned about educational issues such as *over-reliance* on chatbots. There are two main implications of an excessive dependence on AI tools perceived by the students. The first is that such reliance might hinder their exploratory learning. The second is the risk of copying assignments directly without any personal engagement or agency, which could lead to plagiarism and academic misconduct. Over-reliance on AI tools can act as a crutch, reducing the students' independent thinking and creativity.

The aspect of *academic misconduct* is a limitation on the progress of AI tools development, but it is also a challenge present in academia for a long time and as a society. Thus, we need to create new perspectives about teaching and learning. *Academic misconduct, data privacy, and over-reliance on chatbots* are limitations and challenges, this could hinder the personal development of the students but also of our society in general. Limitations about privacy are present in our society and it becomes much more serious when all our devices are connected, and our knowledge is processed by intelligent systems. It is important to be confident about the several limitations of AI tools, but the concern about how it is ethically used might be something to work on to make it more effective and safer for humans. These findings regarding perceived limitations of AI tools on ethical and educational aspects presented by Lee and Maeng (2023) are linked to UNESCO (2021) which highlighted how an overly optimistic perception of AI tools without proper research could lead to negative effects. This remains an ongoing limitation

because AI tools require input from the users to process and generate output. However, since some users are underage students, developers must protect and handle their data with care.

As noted, none of the studies directly examined teachers' perceptions about the limitations of AI tools used in the context of ESL/EFL K-12 classrooms. There are studies about the perception of EFL teachers in higher education (Gao, 2024; Mohamed, 2024), and most limitations are related to academic integrity and data privacy (e.g., Chung & Jeong, 2024). The RAND report (Diliberti et al., 2024) also brings some information about how teachers perceived limitations on using artificial intelligence in K-12 classrooms, but in general, more research is necessary to investigate perceptions about limitations regarding the use of artificial intelligence in EFL K-12 settings.

Research Question 4: Topics Examined in the Research of Using AI Tools

Finally, I wanted to examine what topics were examined in empirical research about using AI tools to teach ELs. Here I identify and explain the key five topics that were explored in 23 research reviewed: (1) learner anxiety, (2) English proficiency and academic achievement, (3) student-centered approach inside and outside the classroom, (4) vocabulary learning and multimodality, and (5) assessment.

Learner Anxiety

The use of AI tools maintained and decreased the level of anxiety of the students. Chen et al. (2022) found that 30 elementary school students in Taiwan who used AI tools, like Dynamic Assessment-based Speech Recognition (DA-SR) and Corrective feedback-based speech recognition (CF-SR) decreased their level of anxiety in English language speaking and improved

their performance. Similarly, Tai and Chen (2023) investigated Google assistant language learning (GALL) and found that the students increased their communicative confidence and reduced their speaking anxiety. On the other hand, Hsu et al. (2024) used AI-supported image recognition, and they found that in vocabulary performance, the experimental group outperformed the control group, but the level of anxiety and self-regulation stayed the same in both groups. This is remarkably interesting because in higher education Çakmak (2022) conducted a study using a generative AI chatbot called *Replika* with 90 EFL students from a state university in Türkiye. The participating students were tested before and after the intervention and they showed increased levels of anxiety after using the chatbot. The author commented that this result could be related to the fact that the *Replika* did not understand the students' input well, so this might contribute to increasing students' anxiety.

English Proficiency and Academic Achievement

The data from five studies (Hsiao & Chang, 2024; Jeon, 2024; Lee & Maeng, 2023; Liu & Chen, 2023; Wang et al., 2022) showed that English language proficiency or academic achievement or performance influence the use of AI tools. For instance, Hsiao and Chang (2024) examined how the level of English proficiency of the students influenced their performance in online learning. They designed an online course for high school students using an AI smart teaching platform *CooC-Cloud* with three AI-powered tools *Linggle Write*, *Linggle Read*, and *Linggle Search.* They found that students with higher English proficiency demonstrated more engagement and better grades at the end of the program. The influence of English proficiency was also observed when AI tools like chatbots were used. Jeon (2024) reported that the affordances of chatbots were perceived differently by thirty-six Korean primary school students. Those learners

with a higher English proficiency perceived AI tools as opportunities, while novice learners perceived it more like a limitation. The students demonstrated difficulties when interacting in conversation with chatbots. Some of these students did not complete their tasks.

In addition to English proficiency, academic achievement or performance seemed to influence their use of AI tools. For instance, Wang et al. (2022) examined how four clusters of 6th-grade EFL students in China used the AI Coach. They created 4 clusters based on academic performance (i.e., C1- effective learners, C2- Passive learners, C3- well-balanced learners, and C4- Inefficient learners). They found different clusters responded to AI Coach differently. For example, students from C4 (ineffective learners) got more stressed than students from C1 when interacting with AI tools. They concluded that not all students can benefit from the use of AI tools, and the effective use of AI tools will depend on the way students use them. Liu and Chen (2023) also revealed similar findings involving seventy-two elementary school students from Taiwan who were divided into two groups of students (i.e., lower-ability and higher-ability) according to their English proficiency. In their study, higher-ability students from the experimental group benefited more than lower-ability students when they used AI-based Object Detection (ODT) to learn vocabulary.

Student-centered approach inside and outside the classroom

AI tools can cover some aspects of formal and informal learning, as mentioned before. Lee et al. (2023) discussed the role of an AI English learning support system, called Learner-Generated Context (LGC) and how LGC-based English language learning experience could impact the dynamics of learning languages. Their findings show that AI-applied systems could assist learners in creating their learning context without an instructor, curriculum, or place.

According to Lee et al. (2023), LGC is a framework used by researchers to understand how technology is used to support learners to autonomously develop their knowledge. The main aspects of LGC in his study are related to the complexity of knowledge, that is, it is not enough to get knowledge only through textbooks and curricula, students need support to explore other options, the second aspect is the development of technology and how it is connected to the world of the new generation, so LGC framework sees the learners as active individuals who get and create their knowledge through interaction with people, technology and learning content autonomously.

In some studies (i.e., Hsiao & Chang, 2024; Lee et al., 2023; Liu & Chen, 2023) included in this review, AI tools were used to support students during and after their instructions. Liu and Chen (2023) conducted their study with 72 elementary school students from a public school located in rural Taiwan. They were randomly assigned to two experimental and control higher-ability and lower-ability groups. Both groups had the pretest that showed they had similar skill levels before the experiment, then they had instructions for target vocabulary lessons together through the Presentation, Practice, Production model (PPP model), and finally, they practiced the activity using AI-based object detection app technology in separate classrooms. The post-test carried out after 11 weeks (about 2 and a half months) suggested that the experimental group outperformed the control group significantly.

In other situations, they were used for online and offline instruction. Hsiao and Chang (2024) developed online courses for high school students in Taiwan to foster their autonomous writing skills outside the classroom. Students received online instructions from the teacher, then they used AI tools to practice their writing offline after class, and finally in class again they were

supposed to deliver an oral presentation about their writing. One of the students from the experiment conducted by Lee et al., (2023) with Korean high school students showed that one of the most important aspects was the possibility of using the AI system to study anywhere at any time. Findings from these studies show that learning languages inside and outside the classroom, with or without the support of instructors, is possible. It indicates that AI tools can facilitate a student-centered approach in ESL/EFL K-12 classrooms.

Among 23 studies reviewed, two studies (Ye et al., (2022, Wang et al. (2023) specifically mentioned how AI tools could be used across in and out-of-classroom. Ye et al., (2022) conducted an experiment with 50 Chinese high school students who used the Chatbot (Microsoft Xiaoying) to practice conversation after their instructions with a teacher. Both groups had access to the same English teacher, textbooks, instruction sheets, and dialog practice material from the chatbot. For the control group, they received printed material to practice with a partner while the experimental group had a smartphone to use the chatbot as a virtual partner for oral English practice every day after the class. After 28 days, the experimental group improved their grammar, their pronunciation accuracy, and increased their conversational skills. The findings indicate that AI tools can be a valuable resource for students who need a virtual partner to practice outside the classroom and improve their autonomous learning while it is also helpful for teachers assigning specific activities during the instructions in the classroom.

Similarly, Wang et al. (2023) mentioned that the AI coach helped 16 sixth-grade primary school students to create a safe environment inside and outside the classroom and generated beneficial effects on learning. More specifically, the AI coach provided personalized guidance, and the students felt more comfortable using it. The AI coach is a virtual agent created with the

appearance of a human to support unlimited English practice and give continuous feedback. Students can interact through websites or apps on their smartphones. These AI coaches can receive written and oral input from the students and give oral and written output. The AI coach normally starts the conversation, when the student speaks the AI coach transcribes the conversation on the screen and highlights what needs correction followed by feedback and encouragement.

Vocabulary learning and multimodality

AI tools can be used to interpret, read, and generate information in different modes (texts, pictures, audio). They can also mimic and generate additional information and give multimodal feedback (Tai & Chen, 2024a). As a result, multimodality (Kress & Van Leeuwen, 2001) has become an emergent reality in the classroom. I have mentioned before that AI tools have been extensively used for vocabulary learning (Jeon, 2023) and pronunciation. It is possible to see how speech recognition, image recognition, and object detection tools can be useful multimodal AI tools for vocabulary learning.

For instance, Kazu and Kuvvetli (2023) investigated if the practice of pronunciation using AI tools would help 56 high school students from Türkiye to have a longer memory in vocabulary acquisition. They used Artificial intelligence-supported speech recognition and found that students who used the tool had memorized the words for a longer time. Similarly, Hsu et al. (2024) conducted an experimental study to investigate how AI image recognition technologies like image-to-text recognition (ITR) along with self-regulated learning (SRL) affected students' vocabulary acquisition. They found that students in the experimental group outperformed those in the control group. In other words, Hsu et al. (2024) suggested that the use of image recognition

technology helped students to support their self-regulating English learning experience, which involved their senses (sight, hearing). They learned how to identify and pronounce the words when looking for objects in the classroom, and consequently answered the task provided easily. This experience of connecting real objects immersed in the students' reality in different modes contributed to their memory retention of English words.

Assessment

A good assessment is a key element for the success of SLA (Brown, 2004; Swain 2005). Various kinds of assessment, such as integrated performance assessment (IPA) (Adair-Hauck et. al., 2013), dynamic assessment (Vygotsky, 1978), and computer-assisted language testing (CALT) (Chapelle and Douglas, 2006) have been tested by educators with AI tools. One study (Chen et al., 2022) focused on examining the use of AI tools for assessment. They designed a dynamic assessment-based speech recognition (called DA-SR) to improve the English-speaking skills of 56 fifth-grade students in China. Dynamic assessment-based speech recognition is a technology based on Vygotsky's social-cultural theory that recognizes, transcribes, and assesses learners' speech in real time.

Through mediation, it provides scaffolding, guidance, and adaptive feedback as needed to support language development. The data suggested that using automatic speech recognition techniques improved students' speaking skills and reduced their learning anxiety. One notable finding is that stress might affect the students' performance on tests. However, in this study, the use of DA-SR designed as a continuous hands-on activity focused on the students' needs correcting them and giving feedback, then students feel less anxious about future testing when

compared to other methods that use corrective feedback. This finding contributes effectively to the view that assessments can be used to instruct the students and not only to test them.

Research Question 5: Recommendations from Existing Research

The are several recommendations from the studies reviewed about ways to use AI tools in ESL/EFL classrooms. The most frequently mentioned recommendation is about *how the AI tools are implemented to support ESL/EFL classroom practice and the students*. Lee and Maeng (2023) suggested that teachers should develop a clear roadmap with guidelines about how students should use chatbots to make them aware of the benefits and constraints of the tools. Specifically, aspects related to privacy policies and critical thinking about the information provided by the chatbot should be considered. It is recommended that students should be more discerned and critical when using AI tools and make decisions based on their needs. The instructor should be an important ally in inspiring and monitoring how supportive the AI Tools would be for the students.

Similarly, Liu and Chen (2023) recommended changes in the design of some AI tools to accommodate the needs of lower-level students. They suggest modification on the chatbot so it could support vocabulary learning of lower-level students. Additionally, Yang et al. (2022) recommended the improvement of the voice recognition capabilities of the chatbot Ellie to improve its ability to understand the students' oral output. Liu et al. (2024) and Woo et al. (2024) suggested improvement in the chatbot's adaptability to accommodate the preferences of the students and provide scaffolding and tailored instruction. This also aligns with Rakhun Kim (2024) who recommended improvements in the accuracy of AI tools to provide better feedback. Finally, Tai and Chen (2024b) recommended simplifying the language used by the chatbot, which according to elementary school students is overly advanced and complex. The suggestions are

using a simple language and incorporating age-appropriate content for elementary school students according to their level of English proficiency and individual preferences.

The second recommendation is about the *learning context* and *types of topics*. In this study, learning context refers to specific conditions that influence the process of learning English such as the characteristics of the participants, and the educational settings. While the types of topics refer to the content of the English instruction, such as describing daily routines, writing stories, and learning vocabulary about fruits and vegetables. Hsu et al. (2024) recommendations are related to the idea of providing an authentic environment where students can learn and bring their knowledge to real life. Ye et al. (2022) discussed how context can help to create a stable and safe environment for practicing the language.

This environment can also work for after-class activities that could be in person or online. According to Hsiao and Chang (2024) it is possible to include theoretical approaches to design synchronous online courses. Similarly, An et al. (2023) recommended policymakers to consider the context where the research was conducted before applying its findings in their context. For example, An et al. (2023) conducted their work in an AI education demonstration area in China where teachers had support from their government and districts, which may not be the reality in other places. So, their recommendation is to use the data theoretically and compare it to other studies about the use of AI tools in education before implementing them. Jeon (2024) believes the types of topics considered for AI use have a key role in supporting the use of AI tools in ESL/EFL classrooms. He stated that affordances of AI tools can change according to the topic which is presented. For instance, students were more willing to communicate with chatbots when they were familiar with the topics of the tasks.

The third and last recommendation from the research reviewed is to encourage students to experiment or use AI tools given their capabilities of revolutionizing the practice of learning languages. Yuan (2023) highlighted that chatbots would facilitate, increase the proficiency of students, foster confidence, and reduce their anxiety. Hsiao and Chang (2024) also stated that AI tools might be an alternative to build the balance between teacher-centered and student-centered teaching and learning approaches. It would reinforce the role of teachers as mentors to support the students in their decision-making.

Chapter 4: Discussion

Reflecting on the four research questions based on the data I have analyzed, I might say I have found the answers for them, and I have raised other questions. These additional questions might be useful for future research into the topic. The first question; RQ1. What AI tools are used to support the learning of the English language inside and outside the ESL/EFL K-12 classroom? I brought to this review 23 studies with all types of AI tools. I divided these AI tools into three groups (chatbots, AI assistants, and other AI tools). When I compared them to the Taxonomy of AIED Systems developed by Holmes and Tuomi (2022), I found that chatbots and AI personal assistants were among the AI tools listed to support students, as well as AI tools like speech recognition, NLP, Assessment and AI-assisted learning apps. This is a particularly important outcome from this review ad for future research. As I mentioned previously, the AIED taxonomy was used to inform the use of AI tools in education, and in this study, the focus was ESL/EFL K-12 Classrooms. So, I might conclude that the taxonomy can still be used as a comparative framework to see how AI tools from a general education context can be used for teaching and learning languages.

To answer the first question (i.e., AI tools used in research), I still needed to check the context of ESL and EFL. In fact, the studies were developed in countries where the EFL is prevalent such as China and South Korea. These findings align with those of Lo et al., (2024) who conducted a study only with ChatGPT in EFL/ESL context and their findings indicated that 48.6% of the studies were conducted in Asia. The most prevalent domain was writing. So, this might be a limitation in answering this question because the study lacks representative of studies

in the ESL context, specifically in countries, such as the USA and the UK. Although the studies used in this work could not indicate to a good extent the AI tools that support the ESL context in the field of education (i.e., Diliberti et al., 2024) indicate that teachers are using AI tools in the ESL context. It makes sense, once these countries have developed many programs for ESL students due to the quantity of immigrants they have received in the last years.

Regarding the context of inside and outside the classrooms, most studies showed experiments where the students interacted with AI tools and the results showed their improvement in the four skills in special speaking skills. This suggests AI tools can be used by students in their autonomous learning. This interaction they have with AI tools generates data that can be used by instructors to give feedback to the students. Certainly, teachers and students should be concerned about privacy, but as the studies demonstrated, some AI tools were developed with the guidance of teachers, who can guide the students to engage in specific task, this guidance might help them to experience a possible data leak.

From the studies reviewed, I perceive that existing AI tools are used by teachers and students in the context of EFL in the classroom. Outside the classroom, students still need to develop their agency about what they want to learn, so they can use AI tools to support them. The classroom brings this aspect of control, so it is easy for the teacher to follow the students' performance while they are in the classroom and the curriculum requirements need to be covered. The question now is how far AI tools can be used to support the learners in these contexts and who is going to decide how they are going to be used.

One aspect that was confusing was how I should label these tools for a better understanding of the readers. ChatGPT is a chatbot with advanced functionalities when compared

to other chatbots. So, in a study about chatbots, it would be confusing to present data from two different chatbots about how students are learning a language. A classification focused on language learning theories might be more interesting. For example. If we consider Input, interaction, and output as a reference for language learning, how do the AI tools get input? Is it through image, audio, text, or number? How can students interact with AI tools? Computers, smartphones, and other tech devices. And how do the AI tools give the output? All these questions might be useful to reflect on the need for more studies to suggest more systematic ways to use AI tools. In most studies the concerns and issues were concentrated on the safety and how AI tools can undermine learning (Lee & Maeng, 2023), as well as the small size of the experiments.

I found the answer to the second question; RQ2. What English language skills (listening, speaking, reading, and writing) in ESL/EFL K-12 classroom? In some studies, the AI tools were used for different English learning skills, in some cases they covered the four English skills, while others were used for just one (i.e., speaking) or two (i.e., reading, writing). One skill that AI tools were most used to cover was speaking. The field of English language learning has been for a long time a field for testing innovative technologies. With these technologies, reading and writing became easier to assess in many countries while listening and speaking were neglected because of the lack of support and advancements.

For example, in standardized tests like TOEFL, students had reading and writing skills accessed in the classroom, but the tests required extra equipment for listening and an evaluator for speaking. In the classroom, it is not different. Sometimes in countries where the students do not have contact with speakers of the target language, they can use technologies like the internet,

social media and others to listen or speak to native speakers. Now they can do it with AI tools. This finding about the skill specifically might be particularly important to cover some gaps in countries where the level of proficiency of teachers and students is not good, and they do not have the opportunity to practice the language.

The use of AI tools from the study can be connected to SLA theories (e.g., input, output, interaction). They were used to support different English skills in the context of EFL/EFL K-12 classrooms. These tools were previously mentioned by Holmes and Tuomi (2022). Most AI tools used in the context of EFL/ESL classrooms were chatbots, they worked better for the aspect of interaction (Jeon, 2024). Students can follow their own pace, and they might have more opportunities to practice when not inside the classroom. According to Kazu and Kuvvetli (2023) "artificial intelligence-supported speech recognition is a learning method that may enhance students' word recall capability much better than the international phonetic alphabet pronunciation teaching method" (p. 488).

I found the answer to the third question; RQ3. How do teachers and students perceive the effectiveness, challenges, and limitations of AI tools in ESL/EFL K-12 classroom settings? The data showed that teachers were concerned when the administrators do not give them the support and the correct training to use AI tools in EFL/ESL classrooms. The data also showed that in some countries there is support not only for the use of AI tools but in creating an environment specifically for this purpose. The continuous discussions about safety and regulations to start using innovative technologies in K-12 classroom are positive. But the data showed that they are already using these tools; they want to ask permission to use what has already been used.

The perceptions of teachers and students in general are overly optimistic. And these perceptions were affected by the level of contact they had with AI tools. One way that might be helpful for students and teachers is to start experimenting with these tools by testing them continuously and implementing them according to the affordances. Being optimistic about AI tools does not make them good but using them and learning instruments to solve real problems in the classroom might be. Sometimes, teachers are very motivated, but without support they will not start something without testing it before by themselves. The data showed teachers have used AI tools for class planning, and assessment, but in some cases, they are not ready to effectively use them to support the students. On the other hand, the students use AI tools in their homes, but they also do not know how much of this use can facilitate or hinder their development and learning at school.

I found the answer to the fourth research question; RQ4. What are topics regarding the use of AI tools in ESL/EFL K-12 from the studies reviewed? The data showed a concentration of AI tools used for supporting students through speaking and listening skills. Generally, topics related to these two skills are important for the development of the field and to give English learners from non-English speaking countries opportunities to practice. The studies reviewed covered the following topics: (1) learner anxiety, (2) English proficiency and academic achievement, (3) student-centered approach inside and outside the classroom, (4) vocabulary learning and multimodality, and (5) assessment.

The discussion about anxiety is recurrent in SLA, for instance, Krashen (1985) highlighted how high anxiety negatively influences the comprehensible input of English learners (ELs). In this work it is possible to see both perspectives about anxiety. There are students who experienced

AI tools that did not understand their speech well and this increased their anxiety (Jeon, 2024), while other students practiced their speech with AI tools that reduced their anxiety (Tai and Chen, 2023). In addition, the data suggests that English proficiency predicts academic achievement. For instance, students with lower level of English proficiency felt more stressed when using AI tools and consequently they had a lower academic achievement. It is possible to infer these two topics are related. It is observed in the studies that students want to learn but they need support.

This goes to the third topic that is student-centered approach. Sometimes, students do not think like teachers do, but they agree about the skills they are supposed to learn and develop. This agreement between students and teachers is a good match to develop this kind of environment with AI tools. However, there is a question. Who wants to bring AI tools to the classroom? Do they want students to start learning how to use them? Lee et al. (2023) discussed LGC based learning and how to use Learner Generated Context (LGC) framework to support students. He describes the context of South Korea where students have good digital literacy, but due their dependence of high-stake testing do not have opportunity to learn what they want, in this case Lee et al. (2023) suggests that Korean students would be willing to participate of an English class where they could find learning opportunities.

The fourth topic, vocabulary learning, and multimodality brought an element that helped increase the participation of the students when they were using AI tools. The data showed that students could engage more and contribute to the process because the AI tools brought some ents they were not familiar with technologies like speech recognition and image recognition are different, but were used for vocabulary acquisition and pronunciation, they brought image, sound, words. Teachers sometimes ask themselves how a specific tool can be important to cover the

needs of students and build a more equitable environment. And I believe that working with a topic such important as vocabulary but in different mediums has helped teachers and students to perceive empowerment of multimodality through AI tools as a common practice for the students. The data from all 23 studies suggest this multimodal approach creates an environment where the students can be differentiated by their skills and receive the support they need.

The last topic ‘assessment’ also contributed to give the answer for the questions teachers normally ask about this integration. Nowadays, it is challenging to assess the students and help them to achieve because there are two models. The first is how to use assessment to differentiate learning to support the students and help them to go beyond their skills, it is the plus 1 (i+1). There are public schools, private schools, schools for gifted students, and others for special education. It looks that schools do not talk to each other, and universities do not talk to the schools. But studies have shown that AI tools might be used to cover this gap of assessment. This aspect of Dynamic assessment might not cover this huge gap in the short term. But using the information provided by AI tools appropriately can be a way to support students from different types of schools. The answer comes from everyone. I think a continuous evaluation about what the students think about what they have received. What teachers think about what they have, and what administrators think about what they have managed.

Finally, regarding the fifth research question (i.e., recommendations by research reviewed), the research suggests that even when using AI tools, teachers should consider the level of the students (Jeon, 2023). This recommendation is one of the most relevant for the field because it affects the way AI tools are developed and the users directly. The AI tools used were in most cases chatbots, they are also personal and teaching assistants. These tools can adapt to

distinct levels of proficiency, but the results showed that in a formal environment students might not feel confutable with them when they do not know how to use them. These results showed the importance of preparing students on how to use the technology and checking how much the tools can cover their needs.

Chapter 5: Conclusion and Future Research

The aim of this narrative research synthesis was to review current empirical studies in the context of ESL/EFL K-12 classroom regarding the use of AI tools to support EL learning and teaching. After conducting an electronic search in two databases (ERIC, PsycINFO), I selected 23 studies. The most common AI tools from the studies were chatbots, and AI personal assistants, which were previously mentioned in the taxonomy of AIED systems by Holmes and Tuomi (2022). The data demonstrated that AI tools were used by students, and teachers covering the four English skills (reading, writing, speaking, and reading) according to the SLA framework (input, output, and interaction) proposed by Krashen, Long, and Swain (1985). The results also showed the relationships between AI tools and anxiety, environment, vocabulary learning, multimodality, and English level. Speaking skills were the most supported skill. Teachers were also concerned about privacy, academic misconduct, the extent of support they have received from administrators, and how AI tools should be regulated for classroom use.

To get these results, I analyzed the studies and categorized the AI tools into three groups (chatbot, AI personal assistant, and others). To answer the RQ1, I listed all AI tools by name and type, and then I found that chatbots represented 43.5%, AI personal assistants represented 17.4%, and Other AI tools represented 39.1% of the AI tools used in the studies. The data showed that these AI tools were used for specific English skills, inside and outside the classroom. To answer RQ2, the data showed that specific English skills were taught with the use or support of AI tools in order to support teachers and English learners. Importantly, speaking was the most popular

skill focused on the studies. Part of the interactions involved pronunciation and vocabulary acquisition including feedback and recast.

To answer RQ3, I analyzed studies about the perceptions of teachers and students, and they showed that the teachers were motivated to use AI tools in their practice in different environments. The data from some studies, (e.g. Ericsson et al., 2023) showed that students had fun using AI tools, but they were also frustrated when the AI could not understand them. They also described them as effective and mentioned concerns about data privacy and better development of AI tools to support students with different levels.

To answer RQ 4, I examined topics examined in the 23 empirical studies reviewed. I identified: (1) learner anxiety, (2) English proficiency and academic achievement, (3) student-centered approach inside and outside the classroom, (4) vocabulary learning and multimodality, and (5) assessment. These five topics could be summarized in just one topic, which is the student-centered approach. Studies showed how anxiety was positive and negative according to the student's English level of proficiency. They also showed the focus on vocabulary and its connection to the four English skills. Topics about learning and assessment were considered. It looks like the topics about AI tools from the studies reviewed are strongly connected to the topics of SLA.

To answer RQ5. I focused on the recommendations of existing research about the use of AI tools in ESL/EFL K-12 classrooms. The data from the studies showed that there was a focus on improving the quality of AI tools, specifically, on how much they understand the users and the second was related to the support teachers would give students about the best ways to use AI tools. In addition, administrators also would have an important role in this aspect of providing

support. Another important recommendation was related to the safety of the students. The way companies should take care of the data of young people.

A further recommendation is to conduct more research on the use of AI tools in multimodality and multiliteracy contexts. Some AI tools from the studies fostered multimodal input, built interaction, and gave the output according to the needs of the learners. It could be image, text sound. This interaction can be covered in different levels of English proficiency and environments such as online, offline, and in-person classrooms. Given that, it will be interesting to examine the extent of support these AI tools will students.

One important finding was the fact that 21 studies (out of 23) were conducted in Asian countries, specifically, China, Taiwan, Hong Kong, and South Korea. This finding about the context can suggest a future direction. For future research, it will be worth exploring diverse environments to understand how they influenced the results of using AI tools in EFL K-12 classrooms. The educational system in different countries and cultures is not as simple as we might picture it. It requires constant evaluation and changes over time. The AI tools used in the process showed that they are used as mediums and even applications with a prominent level of support. According to the recommendation of UNESCO (2021) and Holmes and Tuomi (2022), AI tools should not be used without human supervision if the leaders intend to develop humans in all skills.

The data showed that beyond the four skills, AI tools need more development and research about the novelty effect, especially in ESL K-12 classrooms. Some AI tools were efficient for speaking skills in Asian locations, but they still need to be tested in other countries to measure their efficacy and if there might be any relation to cultural influence. Also, in some experiments,

the authors did not include the aspects of accessibility, but they only tested the AI tools' capabilities. To support teachers, they can be used for planning, instruction, and assessment.

The data from the studies showed how motivated teachers and students feel about using AI tools. The role of the instructor in the process is critical in effectively using AI tools in teaching English. The data showed that AI tools are useful for different tasks in different contexts. However, most experiments were conducted with a human teacher or instructor in the classroom. So, they showed the importance of having someone working as a mentor to support the learners in the process. In some studies, students felt more comfortable using AI tools, than in a regular classroom, but it is necessary more data to understand if this influence came from the relationship between teacher-machine-student or from the aspect of independence students might feel when working by themselves.

In addition, most studies covered aspects of safety, privacy, and honesty. But in most cases, there are still concerns, which indicate that all students should discuss the efficacy of the tools and how safe they are for the users. In higher education this aspect of honesty has been discussed, but what might affect the perceptions of AI tools efficacy in K-12 education is related to how much they can be effective while protecting young students. Parents trust schools to build a safe environment for their children, and this is highly related to their safety and privacy. If for any reason trust is broken, the system may collapse, and parents will see AI tools as a threat to their children instead of a tool that might help them to achieve their goals.

Finally, it is necessary to have more studies with larger samples to build a framework for training teachers on how to use AI tools to support their practice in the classroom. Bearing in mind that it is important to develop critical thinking about the reason behind the use of AI tools

beyond supporting hard skills. Students learn languages at school to better learn other subjects, to interact with other people, and to look for answers on how to make the world a better place for everyone. This should always be the main goal for the educational system when thinking about finding technologies to support teachers and students

References

References marked with an asterisk indicate studies included in the narrative synthesis.

Adair-Hauck, B., Glisan, E. W., & Troyan, F. J. (2013). *Implementing Integrated Performance Assessment*. American Council on the Teaching of Foreign Languages. 1001 North Fairfax Street Suite 200, Alexandria, VA 22314.

Alhalangy, A. G. I., & AbdAlgane, M. (2023). Exploring the Impact of AI on The EFL Context: A Case Study of Saudi Universities. *Journal of Intercultural Communication*, *23*(2), 41–49. https://doi.org/10.36923/jicc.v23i2.125

Alshumaimeri, Y. A., & Alshememry, A. K. (2024). The Extent of AI Applications in EFL Learning and Teaching. *IEEE Transactions on Learning Technologies*, *17*, 653–663. IEEE Transactions on Learning Technologies. https://doi.org/10.1109/TLT.2023.3322128

*An, X., Chai, C. S., Li, Y., Zhou, Y., Shen, X., Zheng, C., & Chen, M. (2023). Modeling English Teachers' Behavioral Intention to Use Artificial Intelligence in Middle Schools. *Education and Information Technologies*, *28*(5), 5187–5208. eric. https://doi.org/10.1007/s10639-022-11286-z

Bahrani, T., Sim, T. S., & Nekoueizadeh, M. (2014). Second Language Acquisition in Informal Setting. *Theory and Practice in Language Studies*, *4*(8), 1714–1723. https://doi.org/10.4304/tpls.4.8.1714-1723

Barrot, J. S. (2023). Using ChatGPT for second language writing: Pitfalls and potentials. *Assessing Writing*, *57*, 100745. https://doi.org/10.1016/j.asw.2023.100745

Benson, P. (2011). Language Learning and Teaching Beyond the Classroom: An Introduction to the Field. In P. Benson & H. Reinders (Eds.), *Beyond the Language Classroom* (pp. 7–16). Palgrave Macmillan UK. https://doi.org/10.1057/9780230306790_2

Bicknell, K., Brust, C., & Settles, B. (2023). How Duolingo's AI Learns what you Need to Learn: The language-learning app tries to emulate a great human tutor. *IEEE Spectrum*, *60*(3), 28–33. IEEE Spectrum. https://doi.org/10.1109/MSPEC.2023.10061631

Brown, H. D. (2004). *Language assessment principles and classroom practices*. White Plains, NY: Pearson Education.

Burkhard, M. (2022). Student Perceptions of AI-Powered Writing Tools: Towards Individualized Teaching Strategies. In *International Association for Development of the Information Society*. International Association for the Development of the Information Society. https://eric.ed.gov/?id=ED626893

Çelik, F., & Yangin Ersanli, C. (2022). The Use of Augmented Reality in a Gamified CLIL Lesson and Students' Achievements and Attitudes: A Quasi-Experimental Study. *Smart Learning Environments*, *9*. eric. https://doi.org/10.1186/s40561-022-00211-z

Chapelle, C. A., & Douglas, D. (2006). *Assessing Language through Computer Technology*. Cambridge University Press.

*Chen, C.-H., Koong, C.-S., & Liao, C. (2022). Influences of integrating dynamic assessment into a speech recognition learning design to support students' English speaking skills, learning anxiety and cognitive load. *Journal of Educational Technology & Society*, *25*(1), 1–14. https://proxy.lib.ohio-state.edu/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=psyh&AN=2022-66872-001&site=ehost-live

Chen, M.-P., Wang, L.-C., Zou, D., Lin, S.-Y., Xie, H., & Tsai, C.-C. (2022). Effects of Captions and English Proficiency on Learning Effectiveness, Motivation and Attitude in Augmented-Reality-Enhanced Theme-Based Contextualized EFL Learning. *Computer Assisted Language Learning*, *35*(3), 381–411. eric. https://doi.org/10.1080/09588221.2019.1704787

Chung, J. Y., & Jeong, S.-H. (2024). Exploring the perceptions of Chinese pre-service teachers on the integration of generative AI in English language teaching: Benefits, challenges, and educational implications. *Online Journal of Communication and Media Technologies*, *14*(4), e202457. https://doi.org/10.30935/ojcmt/15266

Darda, P., Om Jee Gupta, & Susheel Yadav. (2024). Metamorphosing Traditional Pedagogy: Examining the Transcendent Influence of Alexa in Catalyzing Educational Paradigm Shifts within Rural Indian Communities. *International Journal of Educational Management*, *38*(3), 605–621. eric. https://doi.org/10.1108/IJEM-07-2023-0347

Diliberti, M. K., Schwartz, H. L., Doan, S., Shapiro, A., Rainey, L. R., & Lake, R. J. (2024). *Using Artificial Intelligence Tools in K-12 Classrooms*.

Ellis, R. (2021). A short history of SLA: Where have we come from and where are we going? *Language Teaching*, *54*(2), 190–205. https://doi.org/10.1017/S0261444820000038

*Ericsson, E., Sofkova Hashemi, S., & Lundin, J. (2023). Fun and Frustrating: Students' Perspectives on Practising Speaking English with Virtual Humans. *Cogent Education*, *10*(1). eric. https://doi.org/10.1080/2331186X.2023.2170088

Gao, Y. (2024). Exploring EFL university teachers' beliefs in integrating ChatGPT and other large language models in language education: A study in China. *Asia Pacific Journal of Education, 44*(1), 29-44. https://doi-org.proxy.lib.ohio-state.edu/10.1080/02188791.2024.2305173

Gass, S. M., & Mackey, A. (2014). Input, interaction, and output in second language acquisition. In *Theories in second language acquisition* (pp. 194-220). Routledge.

Hattie, J. (2010). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement* (Reprinted). Routledge.

Holmes, W., & Tuomi, I. (2022). State of the art and practice in AI in education. *European Journal of Education*, *57*(4), 542–570. https://doi.org/10.1111/ejed.12533

Howe, K., & Eisenhart, M. (1990). Standards for qualitative (and quantitative) research: A prolegomenon. Educational Researcher, 19(4), 2-9.

*Hsiao, J.-C., & Chang, J. S. (2024). Enhancing EFL reading and writing through AI-powered tools: Design, implementation, and evaluation of an online course. *Interactive Learning Environments*, *32*(9), 4934–4949. psyh. https://doi.org/10.1080/10494820.2023.2207187

*Hsu, T. C., Chang, C., & Jen, T. H. (2023). Artificial Intelligence image recognition using self-regulation learning strategies: effects on vocabulary acquisition, learning anxiety, and learning behaviours of English language learners. *Interactive Learning Environments*, *32*(6), 3060–3078. https://doi.org/10.1080/10494820.2023.2165508

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

Hui-Chin Yeh. (2024). Revolutionizing Language Learning: Integrating Generative AI for Enhanced Language Proficiency. *Educational Technology & Society*, *27*(3), 335–353. eric. https://doi.org/10.30191/ETS.202407_27(3).TP01

*Jeon, J. (2023). Chatbot-Assisted Dynamic Assessment (CA-DA) for L2 Vocabulary Learning and Diagnosis. *Computer Assisted Language Learning*, *36*(7), 1338–1364. eric. https://doi.org/10.1080/09588221.2021.1987272

*Jeon, J. (2024). Exploring AI Chatbot Affordances in the EFL Classroom: Young Learners' Experiences and Perspectives. *Computer Assisted Language Learning*, *37*(1–2), 1–26. eric. https://doi.org/10.1080/09588221.2021.2021241

Jiang, M. Y.-C., Jong, M. S.-Y., Lau, W. W.-F., Chai, C.-S., & Wu, N. (2023). Exploring the Effects of Automatic Speech Recognition Technology on Oral Accuracy and Fluency in a Flipped Classroom. *Journal of Computer Assisted Learning*, *39*(1), 125–140. eric. https://doi.org/10.1111/jcal.12732

*Lee, J. E., & Maeng, U. (2023). Perceptions of High School Students on AI Chatbots Use in English Learning: Benefits, Concerns, and Ethical Consideration. *Journal of Pan-Pacific Association of Applied Linguistics*, *27*(2), 53–72. eric. https://doi.org/10.25256/PAAL.27.2.4

* Kazu, I. Y., & Kuvvetli, M. (2023). The Influence of Pronunciation Education via Artificial Intelligence Technology on Vocabulary Acquisition in Learning English. *International Journal of Psychology and Educational Studies*, *10*(2), 480–493. eric. https://proxy.lib.ohio-state.edu/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=eric&AN=EJ1382350&site=ehost-live

Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., … Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences*, *103*, 102274. https://doi.org/10.1016/j.lindif.2023.102274

Konyrova, L. (2024). The Evolution of Language Learning: Exploring AI's Impact on Teaching English as a Second Language. *Eurasian Science Review an International Peer-Reviewed Multidisciplinary Journal*, *2*(2), Article 2. https://doi.org/10.63034/esr-42

Krashen, S. (1981). Second language acquisition and second language learning. Oxford: Pergamon Press

Krashen, S. (1985). The Input Hypothesis: Issues and Implications. Harlow: Longman.

Krashen, S. (1989). We Acquire Vocabulary and Spelling by Reading: Additional Evidence for the Input Hypothesis. The Modern Language Journal, 73, 440-464. http://dx.doi.org/10.1111/j.1540-4781.1989.tb05325.x

Kress, G., & Van Leeuwen, T. (2001). Multimodal discourse: The modes and media of contemporary communication. London: Arnold Publishers.

Larsen-Freeman, D., & Anderson, M. (2011). *Techniques & Principles in Language Teaching* (3rd ed.). Oxford University Press.

Long, M. (1996). The role of the linguistic environment in second language acquisition. In W. C. Ritchie & T. K. Bhatia (Eds.). Handbook of second language acquisition (pp. 413-468). New York: Academic Press.

*Lee, D., Kim, H., & Sung, S.-H. (2023). Development Research on an AI English Learning Support System to Facilitate Learner-Generated-Context-Based Learning. *Educational Technology Research and Development*, *71*(2), 629–666. eric. https://doi.org/10.1007/s11423-022-10172-2

Lightbown, P., & Spada, N. (1990). Focus on form and corrective feedback in communicative language teaching: Effects on second language learning. Studies in Second Language Acquisition, l2, 429448. http://dx.doi.org/10.1017/S0272263100009517

Lightbown, P. M., & Spada, N. (2013). *How languages are learned* (4th ed.). Oxford, UK: Oxford University Press. (n.d.).

*Liu, C.-C., Chen, W.-J., Lo, F., Chang, C.-H., & Lin, H.-M. (2024). Teachable Q&A agent: The effect of chatbot training by students on reading interest and engagement. *Journal of Educational Computing Research*, *62*(4), 1122–1154. psyh. https://doi.org/10.1177/07356331241236467

Liu, G. L., Darvin, R., & Ma, C. (2024). Exploring AI-mediated informal digital learning of English (AI-IDLE): a mixed-method investigation of Chinese EFL learners' AI adoption and experiences. *Computer Assisted Language Learning*, 1–29. https://doi.org/10.1080/09588221.2024.2310288

*Liu, P.-L., & Chen, C.-J. (2023). Using an AI-based object detection translation application for English vocabulary learning. *Journal of Educational Technology & Society*, *26*(3), 5–20. psyh. https://proxy.lib.ohio-

state.edu/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=psyh&AN=2024-07480-001&site=ehost-live

Lo, C. K., Yu, P. L. H., Xu, S., Ng, D. T. K., & Jong, M. S. (2024). Exploring the Application of ChatGPT in ESL/EFL Education and Related Research Issues: A Systematic Review of Empirical Studies. *Smart Learning Environments*, *11*. eric. https://doi.org/10.1186/s40561-024-00342-5

Long, M. H. (1985). lnput and second language acquisition theory. In S. Gass & C. Madden (Eds.), *Input in second language acquisition* (pp. 377-393). Rowley, MA; Newbury House.

Long, M. H. (1996). The role of the linguistic environment in second language acquisition. In W. C. Ritchie, & T. K. Bhatia (Eds.), Handbook of second language acquisition (pp. 413-468). New York: Academic Press.

Long, M. M. (2017). Instructed second language acquisition (ISLA). *Instructed Second Language Acquisition*, *1*(1), 7–44. https://doi.org/10.1558/isla.33314

Mauldin, M. L. (1994). Chatterbots, TinyMuds, and the Turing Test: Entering the Loebner Prize Competition. Proceedings of the Twelfth National Conference on Artificial Intelligence (Vol. 1, pp. 16–21). AAAI Press.

Mohamed, A.M. Exploring the potential of an AI-based Chatbot (ChatGPT) in enhancing English as a Foreign Language (EFL) teaching: perceptions of EFL Faculty Members. *Educ Inf Technol* 29, 3195–3217 (2024). https://doi-org.proxy.lib.ohio-state.edu/10.1007/s10639-023-11917-z

Mustroph, C., & Steinbock, J. (2024). ChatGPT in foreign language education - friend or foe? A quantitative study on pre-service teachers' beliefs. *Technology in Language Teaching & Learning*, *6*(1), 1133. https://doi.org/10.29140/tltl.v6n1.1133

Nation, I. S. P. (2001). *Learning Vocabulary in Another Language*. Cambridge: Cambridge University Press.

National Institute of Child Health and Human Development (NICHD) (2000). *Report of the National Reading Panel. Teaching Children to Read: An Evidence-Based Assessment of the Scientific Research Literature on Reading and Its Implications for Reading Instruction* (NIH Publication No. 00-4769). Washington, DC: U.S. Government Printing Office. https://www.nichd.nih.gov/sites/default/files/publications/pubs/nrp/Documents/report.pdf

Norris, J. M., & Ortega, L. (2006). Synthesizing research on language learning and teaching (Vol. 13). John Benjamins Publishing.

OpenAI. (2023). *ChatGPT* [Large language model]. OpenAI. https://openai.com/chatgpt

Ortega, L. (2009). *Understanding Second Language Acquisition* (1st ed.). Routledge.

Ortega, L. (2015). Research synthesis. In B. Paltridge & A. Phakiti (Eds.). Research methods in applied linguistics: A practical resource (pp. 225-244). Bloomsbury Publishing.

Ortega, L. (2017). New CALL-SLA Research Interfaces for the 21st Century: Towards Equitable Multilingualism. *CALICO Journal*, *34*(3), 285–316. https://doi.org/10.1558/cj.33855

*Rakhun Kim. (2024). Effects of Learner Uptake Following Automatic Corrective Recast from Artificial Intelligence Chatbots on the Learning of English Caused-Motion Construction. *Language Learning & Technology*, *28*(2), 109–133. eric. https://proxy.lib.ohio-state.edu/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=eric&AN=EJ1428076&site=ehost-live

Radanliev, P. (2024). Artificial intelligence: reflecting on the past and looking towards the next paradigm shift. *Journal of Experimental & Theoretical Artificial Intelligence*, 1–18. https://doi.org/10.1080/0952813X.2024.2323042 eric. https://proxy.lib.ohio-state.edu/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=eric&AN=EJ1428076&site=ehost-live

Rodgers, M., Sowden, A., Petticrew, M., Arai, L., Roberts, H., Britten, N., & Popay, J. (2009). Testing Methodological Guidance on the Conduct of Narrative Synthesis in Systematic Reviews: Effectiveness of Interventions to Promote Smoke Alarm Ownership and Function. *Evaluation*, *15*(1), 49–73. https://doi.org/10.1177/1356389008097871

Shanahan, T. (2000). Research synthesis: Making sense of the accumulation of knowledge in reading. In M. L. Kamil, P. B. Mosenthal, P. D. Pearson, and R. Barr (Eds.), Handbook of reading research (Vol. 3, pp. 209-226). Mahwah, NJ: Erlbaum.

Shafiee Rad, H., & Roohani, A. (2024). Fostering L2 Learners' Pronunciation and Motivation via Affordances of Artificial Intelligence. *Computers in the Schools*, 1–22. https://doi.org/10.1080/07380569.2024.2330427

Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. Gass & C. Madden (Eds.), *Input in second language acquisition* (pp. 235-253). Rowley, MA: Newbury House.

Swain, M. (2000). The output hypothesis and beyond: Mediating acquisition through collaborative dialogue. In J. P. Lantolf (Ed.), Sociocultural theory and second language learning (pp. 97–114). Oxford: Oxford University Press.

Swain, M. (2005). The output hypothesis: Theory and research. In *Handbook of research in second language teaching and learning* (pp. 471-483). Routledge.

Turing, A. M. (1950). I.—COMPUTING MACHINERY AND INTELLIGENCE. *Mind*, *LIX (*236), 433–460. https://doi.org/10.1093/mind/LIX.236.433

*Tai, T.-Y., & Chen, H. H.-J. (2023). The Impact of Google Assistant on Adolescent EFL Learners' Willingness to Communicate. *Interactive Learning Environments*, *31*(3), 1485–1502. eric. https://doi.org/10.1080/10494820.2020.1841801

*Tai, T.-Y., & Chen, H. H.-J. (2024). Improving elementary EFL speaking skills with generative AI chatbots: Exploring individual and paired interactions. *Computers & Education*, *220*, 1–14. psyh. https://doi.org/10.1016/j.compedu.2024.105112

* Tai, T.-Y., & Chen, H. H.-J. (2024). The Impact of Intelligent Personal Assistants on Adolescent EFL Learners' Speaking Proficiency. *Computer Assisted Language Learning*, *37*(5–6), 1224–1251. eric. https://doi.org/10.1080/09588221.2022.2070219

UNESCO. (2021). *AI and education: Guidance for policy-makers*. UNESCO. https://doi.org/10.54675/PCSP7350

UNESCO. (2023). *Global Education Monitoring Report 2023: Technology in education: A tool on whose terms?* (1st ed.). GEM Report UNESCO. https://doi.org/10.54676/UZQV8501

Vygotsky, L. S. (1978). *Mind in Society: Development of Higher Psychological Processes* (M. Cole, V. Jolm-Steiner, S. Scribner, & E. Souberman, Eds.). Harvard University Press. https://doi.org/10.2307/j.ctvjf9vz4

*Wang, X., Liu, Q., Pang, H., Tan, S. C., Lei, J., Wallace, M. P., & Li, L. (2023). What matters in AI-supported learning: A study of human-AI interactions in language learning using cluster analysis and epistemic network analysis. *Computers & Education*, *194*, 1–17. psyh. https://doi.org/10.1016/j.compedu.2022.104703

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

Weizenbaum, J. (1966). ELIZA—a computer program for the study of natural language communication between man and machine. *Communications of the ACM*, *9*(1), 36–45. https://doi.org/10.1145/365153.365168

*Woo, D. J., Guo, K., & Susanto, H. (2024). Exploring efl students' prompt engineering in human–ai story writing: An activity theory perspective. *Interactive Learning Environments*. psyh. https://doi.org/10.1080/10494820.2024.2361381

*Woo, D. J., Wang, Y., Susanto, H., & Guo, K. (2023). Understanding English as a Foreign Language Students' Idea Generation Strategies for Creative Writing with Natural Language Generation Tools. *Journal of Educational Computing Research*, *61*(7), 1464–1482. eric. https://doi.org/10.1177/07356331231175999

*Yang, H., Kim, H., Lee, J. H., & Shin, D. (2022). Implementation of an AI Chatbot as an English Conversation Partner in EFL Speaking Classes. *ReCALL*, *34*(3), 327–343. eric. https://doi.org/10.1017/S0958344022000039

Yi, Y. (2021). *Reconceptualizing the writing practices of multilingual youth: Towards a symbiotic approach to in-and out-of-school writing*. Routledge.

*Ye, Y., Deng, J., Liang, Q., & Liu, X. (2022). Using a smartphone-based chatbot in efl learners' oral tasks. *International Journal of Mobile and Blended Learning*, *14*(1), 1–17. psyh. https://doi.org/10.4018/IJMBL.299405

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

Zhang, X., Sun, J., & Deng, Y. (2023). Design and application of intelligent classroom for English language and literature based on artificial intelligence technology. *Applied Artificial Intelligence*, *37*(1). psyh. https://doi.org/10.1080/08839514.2023.2216051

Zhu, J., Zhang, X., & Li, J. (2024). Using AR Filters in L2 Pronunciation Training: Practice, Perfection, and Willingness to Share. *Computer Assisted Language Learning*, *37*(5–6), 1364–1396. eric. https://doi.org/10.1080/09588221.2022.2080716