Chinese Language Learning Sciences

Yu-Ju Lan
Grace Yue Qi
Dorothy Chun *Editors*

AI-Mediated Language Education in the Metaverse Era

Springer

Chinese Language Learning Sciences

This book series investigates several critical issues embedded in fundamental, technical, and applied research in the field of Chinese as second language (CSL) learning and teaching, including learning mechanism in the brain, technology application for teaching, learning and assessment. The book series discusses these issues from the perspectives of science (evidence-based approach) and technology. The studies in the book series use the methods from the fields of linguistics (such as corpus linguistics and computational linguistics), psychological and behavioural sciences (such as experimental design and statistical analyses), informational technology (such as information retrieval and natural language processing) and brain sciences (such as neuroimaging and neurolinguistics). The book series generally covers three main interdisciplinary themes: (1) fundamental investigation of Chinese as a first or second language acquisition, (2) development in Chinese language learning technology, and (3) applied research on Chinese language education.

More specifically, the book series involves seven research topics:

- language transfer mechanism in Chinese as a second language
- factors of Chinese as a second language acquisition in childhood
- cultural influence on Chinese acquisition
- information technology, corpus
- teaching material design
- teaching strategies and teacher training
- learning models
- assessment methods

All proposals will be sent out for external double-blind review. Review reports will be shared with proposers and their revisions will be further taken into consideration. The completed manuscript will be reviewed by the Series Editors and editorial advisors to ensure the quality of the book and also seek external review in order to ensure quality before formal publication.

Please contact Alice Xie (e-mail: alice.xie@springernature.com) for submitting book proposals for this series.

Yu-Ju Lan · Grace Yue Qi · Dorothy Chun
Editors

AI-Mediated Language Education in the Metaverse Era

Springer

Editors
Yu-Ju Lan
Department of Chinese as a Second Language
National Taiwan Normal University
Taipei, Taiwan

Grace Yue Qi
School of Humanities, Media and Creative Communication
Massey University
North Shore Auckland, New Zealand

Dorothy Chun
Department of Education, The Gevirtz School
University of California
Santa Barbara, CA, USA

ISSN 2520-1719 ISSN 2520-1727 (electronic)
Chinese Language Learning Sciences
ISBN 978-981-95-0244-8 ISBN 978-981-95-0245-5 (eBook)
https://doi.org/10.1007/978-981-95-0245-5

Contents

Chapter 1
Introduction

Yu-Ju Lan, Grace Yue Qi, and Dorothy Chun

1.1 Context and Scope

The emerging ubiquity of Artificial Intelligence (AI), which involves many different technologies that can be used for a myriad of educational purposes, may prove historically to be a turning point in the evolution of the Internet, particularly in the so-called Metaverse Era. The Metaverse is roughly conceived as a shared virtual space or digital ecosystem in which users can interact in a computer-generated environment with digital objects and/or with both human and simulated users, through a variety of virtual technologies, such as AI and Extended Reality (XR). AI technologies include generative AI, e.g., chatbots such as ChatGPT, automatic speech recognition (ASR), automatic writing evaluation (AWE), machine translation, and intelligent systems. XR technologies encompass Virtual Reality (VR), Augmented Reality (AR), and Mixed Reality (MR). These technologies have the potential to have an impact not only on education, but on society as a whole, because they may fundamentally change the way in which we live and work.

Research on educational technology ideally involves interdisciplinary approaches. In designing and undertaking interdisciplinary research, it is imperative for researchers to consider our reflexivity referring to "the sets of dispositions and activities by which researchers locate themselves within the research processes whilst also

Y.-J. Lan (✉)
National Taiwan Normal University, Taipei, Taiwan
e-mail: yujulan@gmail.com

G. Y. Qi
Massey University, Auckland, New Zealand
e-mail: g.qi@massey.ac.nz

D. Chun
University of California, Santa Barbara, USA
e-mail: dchun@education.ucsb.edu

Y. Lan et al. (eds.), *AI-Mediated Language Education in the Metaverse Era*, Chinese Language Learning Sciences, https://doi.org/10.1007/978-981-95-0245-5_1

attending to how their presence, values, beliefs, knowledge and personal and professional histories shape their research spaces, relationships and outcomes" (Consoli & Ganassin, 2023, p. 1). In addition to reflections that often encourage practitioners and researchers to consider certain aspects of research and teaching practices, reflexivity and doing reflexivity entail researchers' in-depth longitudinal moral commitment to knowledge generation and exchange and making an impact that is beyond academia (Consoli & Ganassin, 2023). Exploring the incorporation and integration of ubiquitous technologies and AI in language education, we position ourselves in this volume as language teachers, researchers, and knowledge contributors and seek to encourage respective contributors from academia and industry to not only narrate an action of reflection, but also reflect on themselves as the recipient of their action. With this perspective on reflexivity, we pivot our humanness to value our relationship to data and participants and to be critical and discerning of our identity and self-positioning as we traverse multiple roles over time. This positionality is a unique strength of the book.

In this volume, in addition to explicating the various definitions and conceptions of the Metaverse, AI, and XR, we seek to begin a dialogue among educational researchers, developers, and practitioners regarding the ways in which these technologies can be leveraged for language learning. It is vital to rethink and reconsider human and AI relationships in transforming the way we teach, learn, and live in everyday life. Reinders et al. (2022) among CALL researchers have emphasized the need to conduct more research on language learning in the classroom *and beyond* that makes real connections with what language learners do for their learning today. What is urgently needed is a consolidation and synthesis of theoretical foundations, visions, and the latest research findings and practices on the integration of AI and ubiquitous technologies and language education from diverse perspectives. This volume, comprising contributions of contextualized AI-mediated language learning and teaching, will provide researchers, practitioners, and teacher educators with an opportunity to closely engage in research-informed practices and intellectual dialogues on both potential opportunities and also challenges and pitfalls in the Metaverse Era.

This volume was inspired by the symposium "AI-mediated Language Education in the Metaverse Era" held at National Taiwan Normal University (NTNU), Taiwan, on June 16–17, 2023. The symposium provided an avenue for experts and researchers to explore three key terms and their intersections: the Metaverse, AI, and Language Learning. The rationale for NTNU to host this symposium stemmed from the rapidly growing availability of two groundbreaking advanced technologies, the Metaverse and Generative AI (such as ChatGPT), and their impact on global society. These technologies have had and will continue to have a profound impact on current and future language learning. Therefore, the symposium facilitated a direct dialogue between researchers, educators, and technologists to deepen and broaden the understanding of how AI and immersive technologies can be integrated into language education, paving the way for innovative teaching methods and learning experiences in language education. The discussions at the symposium highlighted not only the potential but also the challenges associated with the application of these technologies. This book

captures the essence of those discussions, offering insights, and reflections from leading voices in the field.

As an example of the rich dialogue, consider the definitions of the key terms used by the participants and how they were relevant to their work. The term Metaverse, conceived by a few contributors, was first explored in science fiction, "Snow Crash" (Stephenson, 1994) and mentioned in Chap. 10 (Aslam & Li); it has been deemed an Internet evolution (Chap. 3, Qi) that denotes the transcendence between the virtual and the real but not limited by space, either physical or virtual (Chap. 2, Lan). A Metaverse can be seen as a space for a range of socializing activities that will "need to combine fantasy scenarios with simulated real-world ventures" (Chap. 11, Godwin-Jones). In exploring the possibility of the (emergence of the) Metaverse Era, some contributors in the volume discussed advanced technologies, such as AI and Generative AI, and their impact on language learning. Chun et al. (Chap. 4) adopted the definition of AI in Hwang et al. 2020 "a machine-based technique which generates algorithm-based predictions, diagnoses, recommendations, and decisions that influence real or virtual worlds." Aslam and Li shared this definition in Chap. 10, stating that AI is fundamental to machine learning and can generate new original content across various media formats. In relation to Generative AI, a rising trend affecting language learning and education in general, AI relies on machine learning, according to Godwin-Jones in Chap. 11, "applied to an immense dataset of text to find patterns and regularities that allow it to predict the next item in a text sequence." In addition to just text generation, Tu et al. (Chap. 9) amended the definition, explaining that Generative AI can also generate "images, audio, and video content" and that different versions of Generative AI have been updating and expanding their functionalities on a daily basis.

1.2 This Volume's Contributions

As our book aims to contribute both new knowledge and practical insights into AI-mediated language education in the Metaverse Era, we address the following question:

In what ways does this book present educational innovations and new perspectives that contribute to the interdisciplinary research informing effective learning? To answer this question, we revisit the CALL literature and recent developments in the AI front for language learning by drawing on the reflexivity lens. Reflexivity and critical reflections are essential for researchers and practitioners in the areas of Applied Linguistics, Languages and Literacies Education, Educational Technology, and Information Technologies as they work toward addressing the tensions and opportunities between research and practice. Chun et al. (2016) provide a helpful set of reflective questions for language teachers when they consider technology incorporation and integration into their teaching practice. These questions are intended to guide teachers to realize desired learning goals rather than put the focus on a selection of particular technology. We believe these questions are still valid in the (emerging)

Metaverse Era, yet, the final one might face some criticism today due to ethical concerns and academic integrity of incorporating Generative AI in the assessment process.

1. What learning goals do I have for my students?
2. What language, culture, and instructional resources do I have available?
3. How can these resources be used and combined most effectively to serve the established learning goals?
4. How will I assess how effective students' use of these resources is in their attainment of the established learning goals?

Kern's (2024, p. 3) article "Twenty-first century technologies and language education: Charting a path forward" published in the *Modern Language Journal* reminds us of an additional set of questions to consider, namely the role of teachers. In his words, "our profession is truly at an inflection point," in that language professionals and researchers currently under the threat of the new wave of the technology revolution must be ever critical in sustaining the fields of CALL, Language and Literacy Education, and Applied Linguistics. The ubiquitous technologies, including digital tutorial platforms, e.g., Duolingo, machine translation platforms, e.g., Google Translate, Generative AI programs, e.g., ChatGPT, Gemini, have rendered language learning to be personal instruction. In addition, discussion around whether language learning offered at the university and in schools is necessary has also contributed to discontinuation of funding and job losses at some universities (Kern, 2024) and even schools globally. However, as Kern (2024) puts it, "technological innovations are always intertwined with issues of culture, values, and ethics." The importance of human language teachers is key to the humanistic value of studying languages (Walkowitz, 2023) and more importantly, we teachers inspire and provide guidance, scaffolding and mentoring to foster learners to analyze, interpret, evaluate, discern, create critical thinking, and reflect as they navigate the socially constructed world (Kern, 2024).

The chapters in this volume all address the above question and, in particular, include the authors' own reflexivity practices and insights, emphasizing the need for human teachers and researchers to interact critically with new technologies and to be ever mindful of the sociocultural aspects of language teaching and learning, as well as the ethical biases and privacy issues that technologies may ignore.

1.3 Organization of the Book

The chapters, excluding the concluding chapter, are grouped into three themes: (1) Language learning in the emerging Metaverse (Chaps. 2 and 3); (2) AI integration with virtual reality for language learning (Chaps. 4, 5); and (3) AI-mediated language learning: Challenges and possibilities (Chap. 6–10). Chapter 11, the concluding chapter, consolidates the key findings of each contribution and presents the potential

way forward for future CALL research and practice in different educational settings and beyond.

For the first theme, "**Language learning in the emerging Metaverse,**" Chap. 2 by Lan provides an overview of the research trends and the cutting-edge technologies (including the Metaverse and AI) currently being used for language learning. The chapter begins by explaining what the Metaverse is and then explores how such immersive technology can enhance students' language learning by increasing their engagement, motivation, and cognitive development through experiential learning. Next, it examines the research trend of how the Metaverse is employed for language learning by discussing six systematic review or meta-analysis studies. It was found that (1) most of the existing literature on the Metaverse for language learning does not report effect sizes, and (2) most Metaverse-related research has focused on skills at the lower end of Bloom's taxonomy, with "remembering" being the most studied. The chapter suggests some directions for future research on Extended Reality (XR) for language learning, including (1) more large-scale and empirical studies, (2) investigating learners with special needs, (3) providing students with more opportunities for collaborative learning and creation, (4) equipping teachers with essential skills to appropriately integrate the advanced technology into their teaching, and (5) careful evaluation of the impact of the Metaverse and AI on language learning. This chapter concludes by discussing the challenges that researchers and educators should consider when using the Metaverse and AI in language learning, especially the classroom ecosystem, accessibility, negative effects, and pedagogical goals.

Complementary to this overview, Chap. 3 presents a humanistic view of language learning in the Metaverse Era by drawing on language learning as investment and sociomaterialist theoretical lenses. Qi in this chapter, "Towards a Framework of Humanising Language Learning in the Emerging Metaverse: Implementation in a Telecollaboration Project for Adult Learner Intercultural Capabilities," proposes a theoretical framework *humanising language learning in the emerging metaverse* to explore language learner identity construction as they traverse real and virtual spaces over time. She defines "the emerging metaverse" as the new Internet wave that enables individual learners to take ownership of learning, interaction, and communication through the languages they use online and offline and not to be strictly attached to the traditional textbooks, physical places, and the target language communities and cultures. In other words, in the emerging Metaverse, language learners are empowered to connect with the world as they constantly traverse between real and virtual worlds. Afforded by ubiquitous technologies, Qi implemented the framework in a telecollaboration project, *Whakawhanaungatanga*, drawing on Māori knowledge, which means establishing, maintaining, and deepening relationships. This ten-week telecollaboration experience enabled learners of Chinese and English from Aotearoa New Zealand and Taiwan to navigate the meaningfulness and impact of Whakawhanungatanga. Learners' ways of living and embodying the language they learn have facilitated them in exploring and embracing new concepts and making meaningful connections with themselves, their project partners, and the world. Their reflective accounts, digital creations, and evidence of communication have revealed the

potential of the framework in fostering language learners' intercultural capability development through telecollaboration.

The second theme "**AI integration with Virtual Reality for language learning**" is explored by Chun et al. in Chap. 4 "AI-mediated High-immersion Virtual Reality for Language Learning," which finds a lack of *both* theoretical models for combining the use of AI and HiVR for language learning *and* educational apps that take full advantage of generative AI built into the apps to simulate authentic conversations. However, progress is being made, in particular with ASR and generative AI tools such as ChatGPT, which can "comprehend" learner input, whether written or spoken, and respond to the input with plausible and appropriate conversational rejoinders. Future challenges include theorizing about how the affordances of new technologies contribute to language learning and determining whether learner-ChatGPT conversations reflect authentic, socioculturally acceptable discourses. As suggested by Godwin-Jones in Chap. 11, consideration should be given to holistic, ecological perspectives, as well as to complexity theory and frameworks such as sociomaterialism and non-Western ontologies, when seeking to understand the complex relationships between humans and their environment, which includes AI-created interlocutors.

A concrete example of how AI has been integrated with virtual reality is presented in Chap. 5 by Grant "A Comparison of Voice-based versus Text-based Communication with Non-character Players during Complex Tasks on the 3D Virtual World Platform of Second Life: Preliminary Student Feedback and Design Implications." Following a brief history of chatbots and Generative AI, along with a short description of 3D multiuser virtual environments, Grant reports on a pilot study of six learners of Chinese as a foreign language who interacted with non-playing characters (NPCs) in a virtual township of Chinese Island in Second Life, in both text-based chat and voice-based chat. Feedback elicited from the participants on perceived and actual differences, and the challenges, benefits, and disadvantages of both forms of communication revealed that voice-based chat with NPCs can increase motivation (fun) for some, is easier than typing for some, and provides opportunities to practice pronunciation and communicate in a manner closer to that of a real conversation. These results and insights can inform future design and implementation changes, for example, incorporating the latest version of ChatGPT for improved automatic speech recognition and conversational abilities.

Theme Three, "**AI-mediated language learning: Challenges and possibilities,**" consists of five chapters. Chapter 6 "Applying AI Technologies for Second Language Learning" by Chen, Hsieh, and Yen highlights the affordances of AI for CALL tools that can be used outside of the classroom for more personalized, targeted learning. The two major areas of focus are speaking and writing tools that have been adapted for use in teaching EFL at a university in northern Taiwan. These tools rely on the growing sophistication of Automatic Speech Recognition (ASR), Speech Synthesis and Text-to-Speech (TTS), Intelligent Personal Assistants (IPA), high-quality chatbots that are powered by Generative AI (e.g., Cool-EBot), and Large Language Models (LLM), such as ChatGPT, that can function as grammar checkers. The use of such tools brings challenges in discerning the correctness and authenticity of information generated

by AI, the misuse or overuse of such tools by learners and teachers alike, and the need for plagiarism guidelines.

Chen and Yeh in Chap. 7 introduce an innovative learning environment, emulating a real-world environment, which combines AI-driven educational robots and Internet of Things IoT-enabled tangible objects, to foster very young children in foreign language learning. The chapter introduces their design principles and guidelines that can be adapted and applied in early language learning contexts, including the focus on the learner-robot, learner-digital content, and learner-physical teaching aid (IoT-enabled tangible objects) interactions. A sample design was illustrated and discussed. The second part of the chapter highlights the building of a multimodal storytelling model that integrates Generative AI technologies, robots, and IoT-enabled tangible objects to assist researchers, educational technologists, and language practitioners in developing contextualized learning activities. The authors conclude the chapter by proposing that their research-informed practical approach to achieving interactive, context-rich language learning in an emulated real-world environment has the potential for language learning for all age groups.

"AI for Language Learning: Opportunities and Challenges" coauthored by Hsiao, Chang, and Chang is Chap. 8 of the volume providing an overview of the technology behind ChatGPT, including BERT (Bidirectional Encoder Representations from Transformers) and T5 (Text-To-Text Transfer Transformer). The chapter further explores the potential of AI tools for language teaching and learning, including ChatGPT, Bing Image Creator, Diffit, and Midjourney. Specific examples are provided, including the prompts used by teachers to help students improve their English skills and how students used AI tools to write and illustrate books. Teachers' and students' perceptions of the advantages and disadvantages of using ChatGPT for teaching and learning, respectively, are analyzed. In concluding this chapter, the authors discuss future research opportunities alongside challenges that might be confronted by professionals and researchers, such as how to strike the delicate balance of using AI tools, yet avoiding students' overdependence on AI and a potential loss of intrinsic learning responsibility.

Chapter 9 entitled "Framework and Potential Application of GAI-based Language Learning Design" by Tu, Liu, and Chen explores the roles of Generative AI as a tool, a tutor, and a tutee. As a tool, four types of Generative AI, including text-based, image-based, audio-based, and video-based Generative AI, can be applied to different learning stages of Bloom's Taxonomy, progressing from remembering, comprehending, applying, to analyzing, evaluating, and creating. As a tutor, Generative AI tools can provide personalized learning paths to meet the specific needs of individual students. In addition, Generative AI tools can take on the role of a peer learner (tutee), acting as a collaborative partner or facilitator during the learning process. Interacting with Generative AI can also enhance students' critical thinking skills and deepen their understanding of the learning process. In addition to exploring the three roles of Generative AI, the authors also point out potential topics for future research, including exploring learners' behaviors regarding the use of Generative AI technologies to support language learning and their impact on higher-order thinking skills, investigating the potential of multimodal Generative AI for language learning,

and evaluating the ethical issues and cultural biases of Generative AI tools in language education.

Chapter 10 by Aslam and Li titled "Revolutionizing Language Learning: The Convergence of AI and the Metaverse in Educational Paradigms" explores the intersection of AI and language learning in the Metaverse Era from an industry perspective. A review of the literature regarding the influence of AI on language education is provided. This leads to the application and technologies evaluation that discusses the gaps between AI technological innovations and their practical applications in language education specifically. Their chapter draws on the academic lens, but more importantly, highlights the industry perspective to initiate the valuable and necessary dialogue between academia and the education technological industry to navigate the practicality of AI and Metaverse integration in language education. The final remarks of the chapter remind us to reflect on the future direction of research and teaching in language education, by (1) addressing critical gaps in AI and Metaverse language learning research, (2) exploring promising AI and Metaverse research areas, (3) exploring emerging technologies in AI and Metaverse language learning, (4) empowering educators for effective AI and Metaverse integration, (5) including ethical considerations and teacher preparedness.

The final chapter by Godwin-Jones "AI and VR Converge: The Future of Language Learning in an Emerging Metaverse" masterfully discusses the promise and potential of integrating multimodal AI into VR/AR to enable contextualized learning in the Metaverse, where the digital world and the real world meet seamlessly, and interactions between and among human learners and AI chatbots can be immersive, expansive, and spontaneous. Yet, the collaborative space for social learning and the ability for individualized learning provided by Generative AI might be tempered by privacy and ethical issues, as well as by the lack of emotional support for learners (and teachers) to develop their humanistic identities.

References

Chun, D., Kern, R., & Smith, B. (2016). Technology in language use, language teaching, and language learning. *The Modern Language Journal, 100*(S1), 64–80. https://doi.org/10.1111/modl.12302

Consoli, S., & Ganassin, S. (2023). Navigating the waters of reflexivity in applied linguistics. In S. Consoli & S. Ganassin (Eds.), *Reflexivity in applied linguistics: Opportunities, challenges, and suggestions* (1st ed., pp. 1–16). Routledge. https://doi.org/10.4324/9781003149408

Kern, R. (2024). Twenty-first century technologies and language education: Charting a path forward. *The Modern Language Journal*. https://doi.org/10.1111/modl.12924

Reinders, H., Lai, C., & Sundqvist, P. (Eds.). (2022). *The Routledge handbook of language learning and teaching beyond the classroom.* Routledge.

Walkowitz, R. L. (2023, September 19). Gutting language departments would be a disaster. *Chronicle of Higher Education.* https://www.chronicle.com/author/rebecca-walkowitz

Yu-Ju Lan is a Distinguished Professor in the Department of Chinese as a Second Language at National Taiwan Normal University. She is currently the Editor-in-Chief of *Educational Technology & Society*, Research Syntheses Editor of *Language Learning & Technology*, and on the editorial board of *Ampersand*. She is the founding president of the Taiwan Pedagogy and Practice in TELL Association. Before becoming a professor at the university, she had been teaching in two different elementary schools in Taiwan for over 20 years. This experience has motivated her to develop different platforms to meet different language learners' multiple needs. Her research interests include technology-enhanced foreign language learning, virtual reality, pedagogical AI agent, generative AI, and online synchronous teacher training. She has published over 50 SSCI journal papers. Lan has proposed the principles of designing tasks and VR contexts based on empirical evidence. As the need for using VR in learning grows, her pioneering works could provide essential implications for academia, education, and industries. For her outstanding research performance, she was awarded the Outstanding Research Award by the Ministry of Science and Technology (MOST), Taiwan, in 2022.

Grace Yue Qi PhD, FHEA, is a Senior Lecturer in the School of Humanities, Media and Creative Communication, Massey University, Aotearoa New Zealand. Her research interests lie in the epistemological and ontological intersections of language, culture and technology, emphasising the humanistic orientation of language education. Her recent projects and publications include language policy and planning in Australasia, plurilingual education, language teacher agency and identity development, and language learner intercultural communicative competences in real and online multimodal environments.

Dorothy Chun is Professor Emerita of Applied Linguistics and Education at the University of California, Santa Barbara and since 2000 the Co-Editor-in-Chief of Language Learning & Technology. Throughout her career, her passion has been researching and exploring how second languages are learned and how technology can be leveraged to enhance the learning of language and culture. Her research areas include L2 phonology and intonation, L2 reading and vocabulary acquisition, computer-assisted language learning (CALL), telecollaboration for intercultural learning, and immersive virtual reality for development of L1 and L2 multiliteracies. She has conducted studies on cognitive processes in learning with multimedia, on whether using visualizations of pitch curves can help L2 speakers of non-tonal languages learn tonal languages and intonation, and on discourse analyses of online intercultural exchanges. In addition to authoring courseware for language and culture acquisition, including apps for Chinese tone pronunciation, she has collaborated on using augmented reality for L2 vocabulary learning and on several virtual reality research projects for language learning.

Chapter 2
Language Learning in the Metaverse: Research Trends, Opportunities, and Challenges

Yu-Ju Lan

Abstract The metaverse is an emerging technology that provides an immersive virtual environment for individuals to interact both with each other and with digital objects. This chapter discusses the following topics related to the metaverse and language learning. First, it describes the metaverse and explores how virtual reality can enhance student engagement, motivation, and cognitive development. Second, it explores research trends in language learning in the metaverse by discussing three systematic reviews and meta-analyses about the utilization of extended reality (XR) for language learning. Third, it suggests the direction for future research on XR for language learning. Fourth, it addresses the challenges that researchers and educators should consider when using the metaverse in language learning, including the classroom ecosystem, accessibility, negative effects, and pedagogical aims. In summary, the aim was to provide insights into the opportunities and challenges when using the metaverse for language education and to suggest future research trends.

Keywords Metaverse · Virtual reality · Second language · Artificial intelligence · Addiction · Artificial intelligence-mediated language learning

2.1 Introduction

The metaverse is an immersive technology blurring reality and virtuality that breaks through the limitations of time and space and provides users with an immersive sense of presence. The metaverse has recently become popular in the technology industry and will undoubtedly have a large impact on different aspects of human society. The metaverse first appeared more than 30 years ago, but has only become a hot topic recently. Stephenson (1992) coined the term "metaverse" in his 1992 science-fiction novel "Snow Crash," which describes several interesting features of the metaverse. First, the metaverse is a virtual-based successor to the internet. Second,

Y.-J. Lan (✉)
National Taiwan Normal University, Taipei, Taiwan (ROC)
e-mail: yujulan@gmail.com

Y. Lan et al. (eds.), *AI-Mediated Language Education in the Metaverse Era*, Chinese Language Learning Sciences, https://doi.org/10.1007/978-981-95-0245-5_2

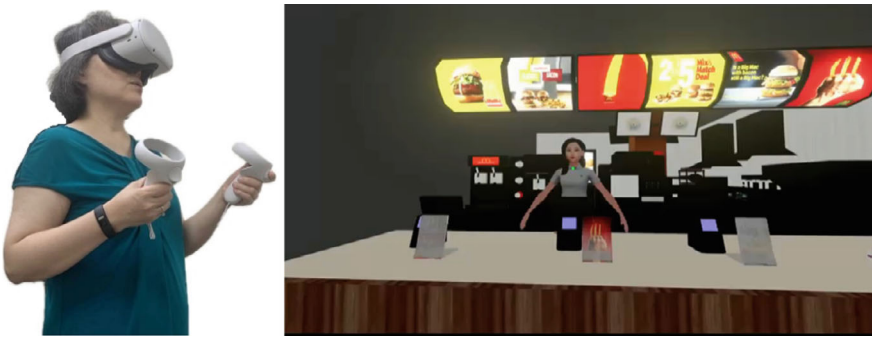


Fig. 2.1 (Left) User wears VR glasses to enter the metaverse, and (right) practices ordering food in a fast-food restaurant

people use customized avatars with goggles and earphones to enter the online world. Third, users can do almost anything in the metaverse that they can in the real world, including exploring, participating in social activities, and conducting commercial trade. The metaverse also provides a way to escape a dystopian reality. Fourth, there is a computer virus in the virtual word that causes metaverse users to suffer brain damage in the real world, which the main character, Hiro Protagonist, attempts to stop.

Most of the features of the metaverse depicted by Stephenson in Snow Crash have become real-life experiences for many users of virtual reality (VR) over the past 20 years. In the example in Fig. 2.1, the user wears VR glasses to practice ordering food in a fast-food restaurant.

The ability of the metaverse to create real-life situations both provides multimodal interactive experiences and allows users to feel immersed in the situation. These features have resulted in this immersive technology being utilized in various fields. Figure 2.2 presents some potential and real applications for the metaverse, such as healthcare (e.g., Chengoden et al., 2023), social networking (e.g., Grant et al., 2024), art creativity (e.g., Yang et al., 2023), simulation (e.g., Rabotin et al., 2023), business (e.g., Popescu et al., 2022), and education (e.g., Hwang & Chien, 2022).

The metaverse can certainly also have a large impact on language learning (Lan, 2021). Lan (2020) highlighted that the metaverse can be segmented into five categories of language education based on differing educational objectives: visual experiences, entertainment, social networking, operation, and creation. But how is the metaverse utilized in language learning, and to what extent does it impact language acquisition? Also, which factors have attracted the interest of researchers, and which emerging trends can be observed in metaverse-assisted language learning? The following sections answer these questions by providing a comprehensive explanation of the metaverse, examining the findings of three systematic reviews on metaverse-based language learning, and providing recommendations for future research.

Fig. 2.2 Potential applications of the metaverse

2.2 What is the Metaverse?

The metaverse emphasizes integrating the virtual and the real. It outlines a blueprint for the future online world, providing people with a living and working environment that appears both real and not limited by space. Ball (2022) suggested that the metaverse is a vast network of three-dimensional (3D) virtual worlds (3DVWs), all interconnected and rendered in real time. It supports an unlimited number of users, each with their own experiences within the world. Moreover, people worldwide can enter these spaces together and interact seamlessly and continuously. Continuous data records are held for the users, including their identities, usage histories, objects owned, and even their communications and transactions.

As depicted in the reality–virtuality continuum proposed by Milgram et al. (1995), which covers the scope between entirely real and entirely virtual environments, the metaverse today is an extended reality (XR), and an immersive technology that encompasses various burgeoning technologies, and holds the promise of advancing interactions in both the virtual and physical realms (Wiles, 2022). Figure 2.3 shows that XR is an umbrella term that encompasses any technology that alters reality by adding any number of digital elements to the physical or real-world environment, thereby blurring the distinction between the physical and digital worlds. This means that XR includes augmented reality (AR), mixed reality (MR), VR, and any other

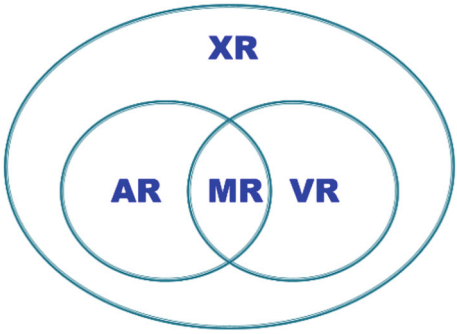


Fig. 2.3 The umbrella of XR

technology that appears at any point along the virtuality continuum, including those that have not yet been developed. The sections below methodically elucidate the three types of immersive technology.

2.2.1 Virtual Reality

VR allows humans to visualize, manipulate, and interact with computers and extremely complex data (Aukstakalnis & Blatner, 1992). It is an artificial environment created using computer hardware and software that is presented to the user to appear and feel like a real environment (Baieier, 1993).

VR can be classified according to the level of immersion or functionality it provides. For example, Robertson et al. (1993) classified it according to level of immersion: immersive and nonimmersive VR. These were renamed by Parong and Mayer (2021) as high- and low-immersive VR, respectively. Immersive VR (Robertson et al., 1993) and high-immersive VR (Parong & Mayer, 2021) therefore mean that the user needs to wear VR glasses and enter the virtual world from a first-person perspective, and hence is isolated from the real world. In nonimmersive VR (Robertson et al., 1993) and low-immersive VR (Parong & Mayer, 2021), they use a mouse or keyboard to control and explore the 3DVW presented on the desktop or laptop screen from a third-person perspective. Regarding the above classification, the participants in the study of Lan et al. (2019) visited a zoo and two restaurants via low-/nonimmersive VR to collect writing ideas, while the platforms introduced in Chun et al. (2022), such as Wander and ImmerseMe, are commercial high-immersive/immersive VR platforms that allow users to wear VR glasses to dynamically experience virtual trips.

Another classification method was proposed by Papagiannidis et al. (2008). They divided VR into two categories based on application functionality: game- or social-based. For example, Lan et al., (2024a, 2024b) created a game-based learning environment in Second Life to help two children with attention deficit hyperactivity disorder to learn Mandarin Chinese, and found very promising effects. Lai and Chen

(2023) also found that learners of English as a foreign language (EFL) who played VR games for their English vocabulary learning outperformed their peers who learned by playing personal computer games.

2.2.2 *Augmented Reality*

AR is a view of the physical world augmented by digital information. It extends the sensory perception of users by utilizing contemporary visualization technology (Buchner & Kerres, 2023). The view is often indirect as it is displayed on a phone or tablet screen. Figure 2.4 presents an AR application that can help elementary-school students in Singapore to learn how to converse orally in Mandarin Chinese (Lan et al., 2024a, 2024b). They were immediately presented with and made able to practice Mandarin with a virtual agent on a mobile device screen after scanning the QR code.

AR provides users with a new environment in which they encounter both virtual and real components (Zhang, 2018), which provides a novel learning experience. This aspect has resulted in AR recently being widely integrated into educational applications such as for language learning. Cai et al. (2022) performed a meta-analysis of language learning that found AR to have a large effect on language gains and a small-to-moderate effect on motivation for learners. Moreover, the positive effects of AR on language learning include that elementary-school students benefit the most in terms of both language gains and motivation.

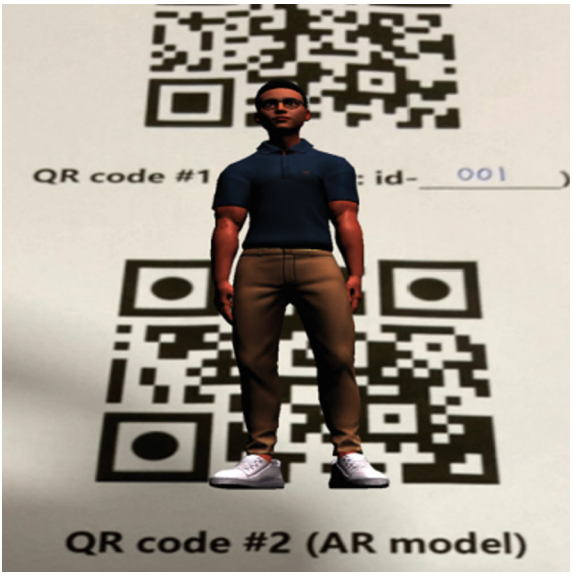


Fig. 2.4 An AR application for learning Mandarin Chinese

2.2.3 *Mixed Reality*

The previous two sections outlined how VR immerses users in an artificial world to create a sense of presence, while AR superimposes virtual objects into the real world. In contrast, MR merges the real and virtual worlds to form a new environment where real-world objects seamlessly coexist with digital ones to provide a real-time interactive experience. MR is an amalgamation of AR and VR, characterized by its real-time interactivity and 3D registration (Yusoff et al., 2011). Although MR is similar to AR, the users have a direct view by wearing MR headsets or glasses as shown in Fig. 2.5. The figure shows two users wearing MR glasses and concurrently observing and manipulating a 3D model of the human body.

The blended space created by MR intentionally closely intertwines the physical and virtual environments. This unique feature of MR means that it has been widely applied across diverse fields, including education (e.g., Torelli et al., 2020), entertainment (e.g., Pratticò & Lamberti, 2020), and medicine (e.g., Lungu et al., 2021). MR also has had a large impact on language learning research, and has been shown to

Fig. 2.5 MR application for observing and manipulating a model of the human body

benefit learners in three ways (Panagiotidis, 2021): reducing their anxiety, improving their cultural awareness and language skills, and improving their general language learning skills.

2.3 Research Trends in Language Learning in the Metaverse

As mentioned above, the metaverse has already been used in language learning. However, there are still various questions that need to be addressed, such as: How is this kind of immersive technology used in language learning? How much influence does the metaverse have on language learning? What factors are more researchers interested in? What are the research trends of language learning in the metaverse? These questions are answered below by discussing six systematic reviews and meta-analyses on using the metaverse for language learning. The systematic reviews are discussed first, followed by the meta-analyses.

The first study is authored by Lin and Lan (2015), and analyzed the research topics, technologies used, language learning environments, sample groups, and methodologies of 29 studies reported on during 2004–2013. The following findings were reported: First, the three most popular research topics were (1) interactive communication; (2) behaviors, affections, and beliefs; and (3) task-based instruction. Higher education was the most common educational setting (25 of the 29 studies), which was much more than that for K-12 (1 for secondary school and 3 for high school). It was clear that foreign language (FL) educators (including those of English, German, and Spanish, but most commonly those of EFL and English as a second language [L2]) were eager to deliver language instruction via VR technologies in higher education over the past 20 years, but nonimmersive/low-immersive VR was mostly used. Moreover, qualitative research was the most widely used design in that period, but with effect sizes being reported for few of the studies. In other words, although some studies used quantitative research, most of the data collection tools were questionnaires; self-reported data were only collected and analyzed in a few relevant studies.

The second study, by Parmaxi (2023), systematically reviewed 26 studies reported on during 2015–2018. Their analysis mostly had the following foci: (a) technologies used, language learning settings, and educational activity duration; and (b) benefits and limitations of using VR as an educational tool in the language classroom. That analysis indicated that most studies adopted Second Life (nonimmersive) (15 of 26 studies) as the learning platform while 5 studies employed OpenSimulator or customized VR; high-immersive VR was not used in any study. Similar to that found in the analysis of Lin and Lan (2015), the studies were most often conducted in higher education settings (12 of 26 studies), followed by primary school settings (6 studies6) and vocational training (5 studies), while only 1 study were conducted in each of secondary and early childhood education. English was still the most common target

language (16 of 26 studies), while Spanish, Chinese, Japanese, and French were uncommon (in 2 or 1 studies only). Language skills were also discussed. It was found that speaking was the skill most often studied (8 of 26 studies), followed by listening (2 studies) and vocabulary (1 study). However, some studies considered some modern skills, such as cultural learning and collaboration. Like in the analysis of Lin and Lan (2015), this systematic analysis notably did not include effect sizes.

The third study, by Parmaxi and Demetriou (2020), was also a systematic review that did not report effect sizes, and analyzed 54 articles published during 2014–2019. However, the main difference was that the technical focus of their analysis was AR, not VR. Regarding student age, although tertiary education still accounted for the largest single proportion of studies (39%), there were also many studies involving elementary-school students (35%). On the contrary, secondary education was relatively uncommon (7%). Notably, preschool children participated in the AR experiment, which was not the case in the systematic review of VR. This might have been due to the devices used in AR-related studies. Most used mobile-based AR devices (63%), which are more popular and less expensive than VR devices. Regarding the target languages, English as an L2 or EFL was the most common (63%), followed by Chinese (14%). It was particularly interesting that sign language was the third most common (7%), which was not found in the previous two VR-related systematic review studies. Another difference between the systematic reviews of AR and VR was the language skills addressed in the analyzed studies: vocabulary (23.9%) was the most common, followed by reading (12.7%), speaking (9.9%), and writing (8.5%).

Based on the above three systematic review studies for which effect sizes were not reported, English as an L2 or EFL attracted the most research attention. The VR research conducted in higher education also accounted for the largest proportion. Regarding elementary and secondary school, AR research was more common than VR research in elementary school, but VR was used more often in middle school than in elementary school. Regarding the devices used, nonimmersive/low-immersive is currently still the main focus for VR research, with high-immersive-related research still being in the minority. AR-related research was mostly based on mobile-based AR. The choice of device was mostly affected by its popularity, ease of use, and cost. This could explain why low-immersive VR and mobile-based AR were the most-used technologies. Regarding language skills, vocabulary acquisition was the focus of most AR research, while speaking ability was the focus for VR research. Language ability may have been related to technology characteristics—because VR is a platform that supports social interaction and both voice and text communication, its application in learning speaking ability is obvious. As for most AR research, it focused on vocabulary acquisition due to the ease of material development.

Now that the three systematic review studies have been discussed, the three meta-analyses are considered below. The first was authored by Wang et al. (2020). It was particularly interesting that effect sizes were reported for very few studies, even though the meta-analysis was conducted in 2019. The number of papers collected from two large databases (Web of Science and ERIC) was still insufficient, and so a dissertation search from OpenDissertations was also included in the target paper

pool, and only 13 papers published during 2008–2019 meeting the analysis criteria to be included in that meta-analysis. The analysis revealed that VR-related studies most often involved college students. However, the results indicated that there was no significant difference in the proportions of students in elementary schools, secondary schools, and colleges, all with large effect sizes ($d = 0.635$, $d = 0.830$, and $d = 0.688$, respectively). This finding indicates the applicability of using 3DVWs for language learning for these three education levels.

Furthermore, the main focus of VR studies for language learning continues to be L2/FL acquisition. English was also the most common target language, grammar and vocabulary were the skills examined most often, and most research was conducted in informal education contexts.

In addition to learning gains, three types of interaction were identified: (1) the interaction between avatars (AA interaction), (2) the interaction between avatar and context (AC interaction), and (3) the interrelation among avatars and context (ACA interaction). Context represents a combination of the objects and environment (Lan et al., 2013). All primary studies adopted the AC interaction type, whereas the AA and ACA types were mostly adopted before 2014. Second Life was the most commonly used platform before 2014, whereas studies after 2014 adopted a larger number of different types of platforms. Both linguistic and affective gains were investigated before 2014, whereas more studies after 2014 examined linguistic gains. Furthermore, low-immersion VR platforms were used the most in meta-analyses despite the increasing popularity of high-immersion VR devices. The findings of the meta-analysis of Wang et al. (2020) included the following aspects. First, 3DVWs that integrate various teaching methods were likely to greatly enhance both the communication skills and linguistic competence of learners. Second, the effective use and promising results suggested a relationship between 3DVWs and learners. In other words, the great affordance of 3DVWs enables learners to adopt them for language learning in nearly all aspects. Third, 3DVWs have been confirmed to impact the cognition and affection of learners, including linguistic and affective gains.

The second meta-analysis was authored by Qiu et al. (2023), who analyzed 14 articles published during 2015–2021. Since there have been few meta-analyses of the use of VR for language learning, and even though the meta-analysis of Qiu et al. only focused on EFL learning, it was still selected and discussed in this chapter. This meta-analysis also confirmed the positive effects of VR on language learning (here is for EFL learning). However, significant differences were found between education levels: a large positive effect for elementary and middle schools, and a small positive affect for colleges, but a negative effect for high schools.

The skills or constructs considered in the analyzed articles indicated that the affect aspect (including anxiety and motivation) was the most common, while English skills were the least common. Moreover, among English skills, vocabulary acquisition and oral expression skills were considered more often than listening, reading, and writing skills.

The last meta-analysis of Cai et al. (2022) focused on AR for language learning, and comprised 21 studies published during 2008–2020. This meta-analysis found that AR application had a large effect on the language gains of learners ($g = 0.93$),

Fig. 2.6 The lower end of Bloom's taxonomy, with the "remember" category at the bottom being the most commonly investigated

but only a moderate effect on their motivation ($g = 0.42$). Regarding education levels, elementary-school students experienced the largest effect on their language gains ($g = 1.23$), followed by students in preschools ($g = 0.81$), secondary schools ($g = 0.8$), and colleges ($g = 0.58$). On the other hand, AR technology had a large impact on the motivation of elementary-school students ($g = 0.72$), but lesser impacts on those in secondary schools ($g = 0.19$) and colleges ($g = 0.1$). English was again the most common target language. Regarding language skills, the application of AR had a large impact on vocabulary learning ($g = 1.16$) and reading ($g = 1.04$), but a rather small one on writing ($g = 0.35$).

The results of the above three meta-analyses were similar to those of the three described systematic review studies: English was the most common target language, and vocabulary was the most common language skill. Elementary-school students also experienced large effects on both their language and affect gains.

More information could be obtained by further synthesizing the above six research syntheses (3 systematic reviews and 3 meta-analyses): (1) effect sizes were not reported for most previous studies described in the literature on XR for language learning, and (2) most XR-related research has focused on abilities at the lower end of Bloom's taxonomy (Wilson, 2016), with the "remember" category at the bottom of that taxonomy being the most commonly investigated (Fig. 2.6). Therefore, there is still plenty of room for improvement despite XR being widely used in language learning for nearly 30 years.

2.4 Suggestions for Future Research on Language Learning in the Metaverse

What will be the next step in research on XR for language learning? Below we list some possible steps for readers to consider.

First, more large-scale and empirical studies are needed to obtain stronger evidence to guide the application of XR in language learning. In addressing the gap identified in the previous section, it became evident that effect sizes have not been reported in most of the existing literature. Consequently, there is a great need for empirical evidence to offer educators and researchers objective guidance in the frontline of policy development and language teaching enhancement.

Second, we need to investigate how XR can benefit learners with special needs in addition to ordinary learners, as the existing literature mostly comprises the latter. However, the specific capabilities of XR, including creating real-life scenarios and supporting multimodal interactions, are particularly beneficial for learners with special needs, such as those with autism spectrum disorder. How to effectively use XR to improve learning for those with special needs is therefore an emergent issue that needs more attention and effort from both researchers and educators.

Third, learners must be provided with more opportunities to interact and collaborate with others to create their own worlds. As described in the previous section, most studies related to XR for language learning have focused on the abilities at the lower end of Bloom's taxonomy. Considering that students need to develop critical competences to ensure their future success, an important research topic is how XR technology should be used to effectively cultivate the critical competences needed by students in the modern world.

Fourth, teachers need to be trained on how to incorporate new technologies into their curriculum and adopt new ways to assess student learning outcomes. XR is a technology that can enhance the interest of students in learning since it attracts their attention. However, students might not focus on what the teacher wants them to learn when receiving rich information through XR, or they may only like new experiences and become less interested in the actual learning content. Teachers must therefore know how to use innovative learning activities to allow students to learn high-level skills from XR experiences, such as problem-solving and working with others. In addition, teachers cannot just use paper-and-pencil tests to assess the learning outcomes of students when attempting to train them in high-level skills. Teachers must therefore understand different assessment techniques to effectively evaluate learning outcomes for students.

Finally, the impact of combining XR and artificial intelligence (AI) on language learning must be evaluated. Numerous platforms have already emerged that seamlessly integrate XR and AI to forge immersive language learning environments. For instance, the MetaClassroom by Cheng (Liu et al., 2023) makes it easy for educators to construct VR settings and seamlessly integrate the popular generative AI platform ChatGPT into an VR environment as objects or teaching agents. The impact of this groundbreaking technology on FL acquisition and the learning methodologies of students merits in-depth exploration and discussion.

2.5 Challenges

What are possible challenges that educators and researchers may face when using the metaverse for language teaching and learning? ChatGPT provides a long list when asked about the challenges of using the metaverse for language learning, such as technical difficulties, limited resources, professional development, accessibility, cost, intellectual property, safety and privacy, pedagogical alignment, lack of evidence, and integration with existing curriculum. Below we focus on the following issues: classroom ecosystem, accessibility, negative effects, and pedagogical aims.

2.5.1 Classroom Ecosystem

The classroom can be viewed as an ecosystem comprising various elements that impact each other, such as teachers, students, learning materials, and settings. Therefore, these elements should be collaborative, interactive, and interdependent. The learning that occurs in the classroom is dynamic and should be adaptive and reciprocal, which makes it a complex process.

What will this ecosystem look like if all students wear VR glasses? In this situation, they are interacting in a VR world that creates a sense of isolation from the real world, including from the real people, materials, and settings in the environment. Users may therefore experience a sense of alienation. How can we maintain the valuable characteristics of classroom learning, such as collaboration, interaction, and interdependence?

2.5.2 Accessibility

The above-mentioned issue could be solved using new VR glasses such as the Apple Vision Pro. These devices reduce the sensation of alienation since users can see their surroundings when wearing them. However, this introduces the new challenge of the VR glasses being very expensive, and so not all educators or learners can afford them. Accessibility would therefore become a new problem for teachers and researchers. Without good instructional design, it is impossible to achieve the desired effect of XR if some students have VR glasses while the rest wait to use them. Additionally, accessibility refers to the training needed for effective use of the VR devices rather than just the technology itself. Teachers need to receive training on effectively integrating VR into their teaching, and students need to follow usage guidelines to better enhance their learning.

2.5.3 Negative Effects

A more serious issue requiring the attention of researchers and educators is the possible mental health-related problems that may arise from inappropriate XR use. Numerous studies have confirmed that the inappropriate use or overuse of technologies can impair the cognitive judgment, emotion, and behavior of the users (e.g., Lavoie et al., 2021; Rodideal, 2018; Siyahhan et al., 2011). Educators and researchers need to pay attention to improving VR use in learning to develop the IQ levels and behaviors of young people and students (e.g., Kaimara et al., 2022), and to avoid the negative effects of VR immersion on the health of learners, including addiction, impaired development, distorted self-image, violent tendencies, and ethical issues. We therefore need to pay attention to both the negative and positive impacts of XR on users, especially young learners. Moreover, VR addiction is one of the most worrying aspects. Research has found that VR addiction may negatively impact cognitive development and behavior in children, seemingly echoing the prediction in the novel Snow Crash that a computer virus will damage the human brain in the real world. Ethical issues also need to be addressed. Predators can use VR chat rooms to target children on popular gaming devices, reminding us of the dangers that young metaverse users may encounter (Deldari et al., 2023; Dorasamy et al., 2021). Children and teenagers need special attention and precautions when making friends and imitating behaviors in virtual universes.

2.5.4 Pedagogical Aims

Finally, it should be considered how XR can fit into modern school courses, indicated by the following common statement of the teachers or researchers to the author: "we have put the materials in the metaverse, but students feel bored and don't know what to do in it." This highlights the main problem: the promising benefits of XR cannot be obtained if we simply move existing learning materials and learners into the metaverse. Instead, we must create a smart and contextualized multimodal learning context where learning is learner-centered, self-directed, and authentic, where learners can gain ownership by creation. This is a new pedagogical paradigm that is needed in the AI-mediated learning era (Fig. 2.7).

The pedagogical aim will be to help each learner become an autonomous, responsible, and smart learning agent. Learners can be self-directed, self-regulated, and self-reflective if they are provided with multifaceted and precise information. Learners can then define their learning goals, conduct learning plans, take actions, direct their learning process, identify the issues they encounter, reflect on the results, and use their experience to refine their plans.

But how can it be confirmed that learners are acting responsibly? This problem becomes even more challenging when generative AI (e.g., ChatGPT) ubiquitously affects our daily lives and education (Fui-Hoon et al., 2023; Okaiyeto et al., 2023).

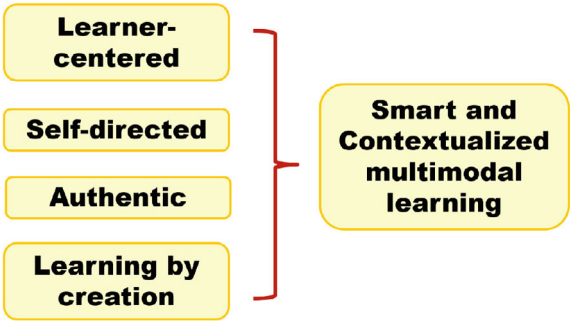


Fig. 2.7 A new pedagogical paradigm

Many educators are concerned that students will become more dependent on AI as they use it more (e.g., Alexander et al., 2023). This makes it easy for them to develop a habitual reliance. The perspectives of educators and researchers indicate that AI may play a role as a learner facilitator, and help students to develop critical thinking skills and to tell right from wrong. Another issue of concern to many researchers is the future of humanity, since we cannot predict how humans will behave and what they will create as they interact with AI. Moreover, open AI not only boosts the positive side of human society, but also strengthens the negative side. For example, WormGPT and FraudGPT can be used by hackers to help them attack and damage the security of our society (Sina, 2023). Even worse, generative AI can be deceptive if the person who controls its training data has bad intentions.

As shown in Fig. 2.8, educators and researchers therefore need a new mindset toward how XR and AI will facilitate smart and contextualized multimodal learning to help learners become autonomous and intelligent agents, and so the following questions should be considered: "How can classroom ecosystems become more balanced, beneficial, and prosocial? And how do students become more responsible and intelligent?".

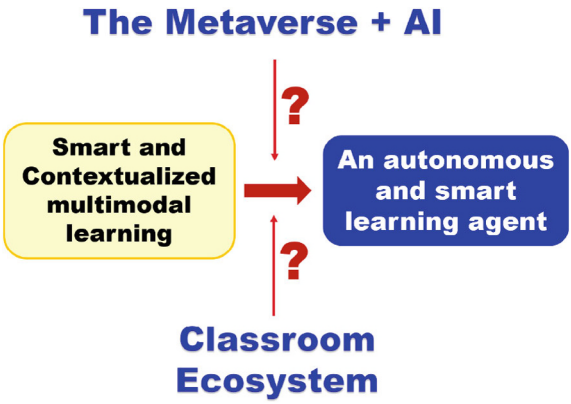


Fig. 2.8 The pedagogical challenge

2.6 Conclusion

When AI and the metaverse become powerful technologies that inevitably affect all aspects of human life, what attitude should educators and researchers adopt toward these technologies? Arguments made by the founder and CEO of Khan Academy Sal Khan (https://youtu.be/hJP5GqnTrNo) may provide us with hints for answering this question. In that TED Talk he shared the opportunities he sees for students and educators to collaborate with AI tools. He suggested that student academic performance will improve significantly if they are provided with personalized one-on-one tutoring. This suggestion could be implemented through carefully designed AI agents (Lan & Chen, 2024) as suggested in the TED Talk by Lee (https://youtu.be/ajGgd9Ld-Wc). Lee suggested that humans and AI can coexist, and that AI could save humanity. He mentioned that AI can help us perform routine jobs and thereby free up considerable time for us to spend with our loved ones. His suggestion is much like flipped learning, in which the learning of lower-level skills or knowledge is rescheduled before students enter the classroom. This allows teachers to use the valuable class time to guide students in developing high-level skills or competences.

Based on the above discussion, should we be against or in favor of using these advanced technologies? This question reminded me of the recently highly discussed movie, Oppenheimer, which portrays the inner conflict of its protagonist. As a scientist, Oppenheimer was eager to see his theories realized, but as a human being, he was once confused by himself becoming the "grim reaper."

In short, are advanced technologies such as the metaverse and AI good or bad? Are they fortunate or unfortunate for mankind? From the perspective of an educator and researcher, I answer this question and summarize this chapter using Fig. 2.9.

The best approach is probably to amplify the impact of the affordances of the metaverse such as embodiment, multimodality, and immersion by integrating an innovative pedagogical approach that involves creation, interaction, and hands-on experiences supported by AI affordances such as personalization, adaptability, and precision. This would allow learners to develop into autonomous, responsible, and smart agents. There is no doubt that we scholars from different disciplines should work together to realize this.

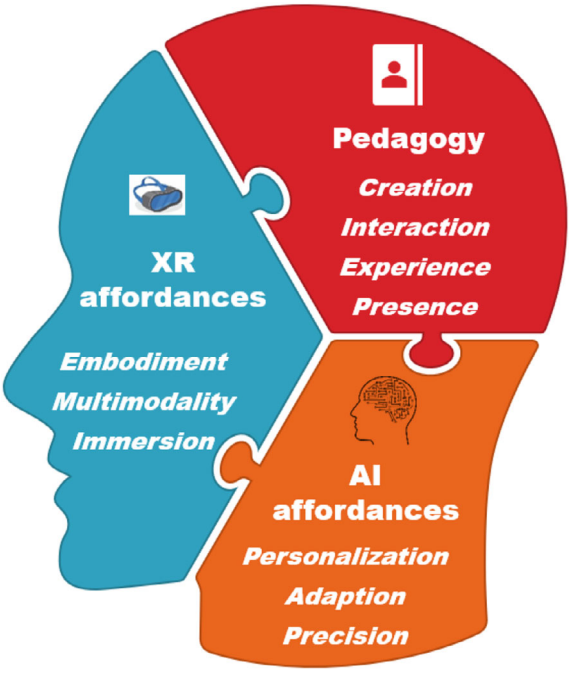


Fig. 2.9 A potential solution to the pedagogical challenge

Acknowledgements The research was supported by the National Science and Technology Council, Taiwan under project numbers MOST 110-2511-H-003 -038 -MY3 and MOST 111-2410-H-003 -006 -MY3.

References

Alexander, K., Savvidou, C., & Alexander, C. (2023). Who wrote this essay? Detecting AI-generated writing in second language education in higher education. *Teaching English with Technology, 23*(2), 25–43.

Aukstakalnis, S., & Blatner, D. (1992). *Silicon mirage; the art and science of virtual reality*. Peachpit Press.

Baieier, K. P. (1993). Virtual reality: Short introduction. https://websites.umich.edu/~vrl/intro/

Ball, M. (2022). *The metaverse: and how it will revolutionize everything*. Liveright Publishing.

Buchner, J., & Kerres, M. (2023). Media comparison studies dominate comparative research on augmented reality in education. *Computers & Education, 195*, Article 104711.

Cai, Y., Pan, Z., & Liu, M. (2022). Augmented reality technology in language learning: A meta-analysis. *Journal of Computer Assisted Learning, 38*(4), 929–945.

Chengoden, R., Victor, N., Huynh-The, T., Yenduri, G., Jhaveri, R. H., Alazab, M., & Gadekallu, T. R. (2023). Metaverse for healthcare: A survey on potential applications, challenges and future directions. *IEEE Access, 11*, 12765–12795. https://doi.org/10.1109/ACCESS.2023.3241628

Chun, D. M., Karimi, H., & Sañosa, D. J. (2022). Traveling by Headset: Immersive VR for Language Learning. *CALICO Journal, 39*(2), 129–149.

Deldari, E., Freed, D., Poveda, J., & Yao, Y. (2023). An Investigation of Teenager Experiences in Social Virtual Reality from Teenagers', Parents', and Bystanders' Perspectives. In *19th Symposium on usable privacy and security (SOUPS 2023*) (pp. 1–17).

Dorasamy, M., Kaliannan, M., Jambulingam, M., Ramadhan, I., & Sivaji, A. (2021). Parents' awareness on online predators: Cyber grooming deterrence. *The Qualitative Report, 26*(11), 3683–3723.

Fui-Hoon Nah, F., Zheng, R., Cai, J., Siau, K., & Chen, L. (2023). Generative AI and ChatGPT: Applications, challenges, and AI-human collaboration. *Journal of Information Technology Case and Application Research, 25*(3), 277–304.

Grant, S., Qi, G. Y., Lan, Y.-L., & Cheng, P.-Y. (2024). Fostering academic citizenship through ubiquitous technologies in an online academic conference: A framework and its implications. *Educational Technology & Society, 27*(1), 18–34. https://doi.org/10.30191/ETS.202401_27(1).RP02

Hwang, G. J., & Chien, S. Y. (2022). Definition, roles, and potential research issues of the metaverse in education: An artificial intelligence perspective. *Computers and Education: Artificial Intelligence, 3*, Article 100082.

Kaimara, P., Oikonomou, A., & Deliyannis, I. (2022). Could virtual reality applications pose real risks to children and adolescents? A systematic review of ethical issues and concerns. *Virtual Reality, 26*(2), 697–735.

Lai, K. W. K., & Chen, H. J. H. (2023). A comparative study on the effects of a VR and PC visual novel game on vocabulary learning. *Computer Assisted Language Learning, 36*(3), 312–345.

Lan, Y.-J. (2020). Immersion, interaction and experience-oriented learning: Bringing virtual reality into FL learning. *Language Learning & Technology, 24*(1), 1–15. http://hdl.handle.net/10125/44704

Lan, Y.-J. (2021). Language learning in virtual reality: Theoretical foundations and empirical practices. In Y.-J. Lan, S. Grant (Eds.) *Contextual language learning. Chinese language learning sciences.* Springer. https://doi.org/10.1007/978-981-16-3416-1_1

Lan, Y.-J., & Chen, N.-S. (2024). Teachers' agency in the era of LLM and generative AI: Designing pedagogical AI agents. *Educational Technology & Society, 27*(1), 1–18. https://doi.org/10.30191/ETS.202401_27(1).PP01

Lan, Y.-J., Foo, K. H. D., & Cheng, P.-Y. (2024). *Enhancing CFL speaking skills with voice-chatbot-embedded virtual avatars: A study on Singaporean elementary students.* AAAL 2024, Houston, Texas, USA.

Lan, Y.-J., Kan, Y.-H., Hsiao, I. Y. T., Yang, S. J. H., & Chang, K.-E. (2013). Designing interaction tasks in Second Life for Chinese as a foreign language learners: A preliminary exploration. *Australasian Journal of Educational Technology, 29*(2), 184–202.

Lan, Y.-J., Lyu, B.-N., & Chin, C.-K. (2019). Does a 3D immersive experience enhance Mandarin writing by CSL students? *Language Learning & Technology, 23*(2), 125–144.

Lan, Y.-J., Shih, M.-F., & Hsiao, Y.-T. (2024b). 3D immersive scaffolding game for enhancing Mandarin learning in children with ADHD. *Educational Technology & Society, 27*(2), 4–24.

Lavoie, R., Main, K., King, C., & King, D. (2021). Virtual experience, real consequences: The potential negative emotional consequences of virtual reality gameplay. *Virtual Reality, 25*, 69–81.

Lin, T.-J., & Lan, Y.-J. (2015). Language learning in virtual reality environments: Past, present, and future. *Educational Technology & Society, 18*(4), 486–497.

Liu, Y. T., Cheng, P. Y., Shih, S. P., & Huang, T. Y. (2023, July). MetaClassroom: A WebXR-Based Hybrid Virtual Reality Classroom. In *2023 IEEE international conference on advanced learning technologies (ICALT)* (pp. 280–281). IEEE.

Lungu, A. J., Swinkels, W., Claesen, L., Tu, P., Egger, J., & Chen, X. (2021). A review on the applications of virtual reality, augmented reality and mixed reality in surgical simulation: An extension to different kinds of surgery. *Expert Review of Medical Devices, 18*(1), 47–62.

Milgram, P., Takemura, H., Utsumi, A., & Kishino, F. (1995). December). Augmented reality: A class of displays on the reality-virtuality continuum. *Telemanipulator and Telepresence Technologies, 2351*, 282–292.

Okaiyeto, S. A., Bai, J., & Xiao, H. (2023). Generative AI in education: To embrace it or not? *International Journal of Agricultural and Biological Engineering, 16*(3), 285–286.

Panagiotidis, P. (2021). Augmented and mixed reality in language learning. *European Journal of Education, 4*(2), 27–43.

Papagiannidis, S., Bourlakis, M., & Li, F. (2008). Making real money in virtual worlds: MMORPGS and emerging business opportunities, challenges and ethical implications in metaverses. *Technological Forecasting and Social Change, 75*(5), 610–622.

Parmaxi, A. (2023). Virtual reality in language learning: A systematic review and implications for research and practice. *Interactive Learning Environments, 31*(1), 172–184.

Parmaxi, A., & Demetriou, A. A. (2020). Augmented reality in language learning: A state-of-the-art review of 2014–2019. *Journal of Computer Assisted Learning, 36*(6), 861–875.

Parong, J., & Mayer, R. E. (2021). Cognitive and affective processes for learning science in immersive virtual reality. *Journal of Computer Assisted Learning, 37*(1), 226–241.

Popescu, G. H., Valaskova, K., & Horak, J. (2022). Augmented reality shopping experiences, retail business analytics, and machine vision algorithms in the virtual economy of the metaverse. *Journal of Self-Governance and Management Economics, 10*(2), 67–81.

Pratticò, F. G., & Lamberti, F. (2020). Mixed-reality robotic games: Design guidelines for effective entertainment with consumer robots. *IEEE Consumer Electronics Magazine, 10*(1), 6–16.

Qiu, X. B., Shan, C., Yao, J., & Fu, Q. K. (2023). The effects of virtual reality on EFL learning: A meta-analysis. *Education and Information Technologies, 1*, 1–27.

Rabotin, A., Glick, Y., Gelman, R., Ketko, I., Taran, B., Fink, N., & Furer, A. (2023). Practicing emergency medicine in the Metaverse: A novel mixed reality casualty care training platform. *Surgical Innovation, 30*(5), 586–594.

Robertson, G. G., Card, S. K., & Mackinlay, J. D. (1993). Three views of virtual reality: Nonimmersive virtual reality. *Computer, 26*(2), 81.

Rodideal, A. A. (2018). Emerging needs for minimizing negative effects of technology overuse among children. *Moldavian Journal for Education and Social Psychology, 2*(1), 1–16.

Sina, A. (2023). Open AI and its Impact on Fraud Detection in Financial Industry. *Journal of Knowledge Learning and Science Technology, 2*(3), 263–281.

Siyahhan, S., Barab, S., & James, C. (2011). Ethics of identity play in virtual spaces. *Journal of Interactive Learning Research, 22*(1), 111–138.

Stephenson, N. (2003). *Snow crash: A novel*. Spectra.

Sutherland, I. E. (1968, December). A head-mounted three dimensional display. In *Proceedings of the December 9–11, 1968, fall joint computer conference, part I* (pp. 757–764).

Torelli, E., El Shemy, I., Silleresi, S., Moskwa, L., Cosentino, G., & Garzotto, F. (2020, March). Do you speak holo? A mixed reality application for foreign language learning in children with language disorders. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)* (pp. 608–609). IEEE.

Wang, C. P., Lan, Y.-J., Tseng, W. T., Lin, Y. T., & Kao, C. L. (2020). On the effects of 3D virtual worlds in language learning- A meta-analysis. *Computer Assisted Language Learning, 33*(8), 891–915. https://doi.org/10.1080/09588221.2019.1598444

Wiles, J. (2022). What is a metaverse? And should you be buying in. *Gartner Insights*. https://www.gartner.com/en/articles/what-is-a-metaverse

Wilson, L. O. (2016). *Anderson and Krathwohl–Bloom's taxonomy revised. Understanding the new version of Bloom's taxonomy.* https://quincycollege.edu/wp-content/uploads/Anderson-and-Krathwohl_Revised-Blooms-Taxonomy.pdf

Yang, X., Cheng, P. Y., Liu, X., & Shih, S. P. (2023). The impact of immersive virtual reality on art education. A study of flow state, cognitive load, brain state, and motivation. *Education and Information Technologies, 1*, 1–20.

Yusoff, R. C. M., Ibrahim, R., Zaman, H. B., & Ahmad, A. (2011). Evaluation of user acceptance of mixed reality technology. *Australasian Journal of Educational Technology, 27*(8), 1369–1387. https://doi.org/10.14742/ajet.899

Zhang, S. (2018). Augmented reality in foreign language education: A review of empirical studies. *Journal of Technology and Chinese Language Teaching, 9*(2), 116.

Yu-Ju Lan is a Distinguished Research Chair Professor in the Department of Chinese as a Second Language at National Taiwan Normal University. She is currently the Editor-in-Chief of *Educational Technology & Society*, Research Syntheses Editor of *Language Learning & Technology*, and on the editorial board of *Ampersand*. She is the founding president of the Taiwan Pedagogy and Practice in TELL Association. Before becoming a professor at the university, she had been teaching in two different elementary schools in Taiwan for over 20 years. This experience has motivated her to develop different platforms to meet different language learners' multiple needs. Her research interests include technology-enhanced foreign language learning, virtual reality, pedagogical AI agent, generative AI, and online synchronous teacher training. She has published over 50 SSCI journal papers. Lan has proposed the principles of designing tasks and VR contexts based on empirical evidence. As the need for using VR in learning grows, her pioneering works could provide essential implications for academia, education, and industries. For her outstanding research performance, she was awarded the Outstanding Research Award by the Ministry of Science and Technology (MOST), Taiwan, in 2022.

Chapter 3
Towards a Framework of Humanising Language Learning in the Emerging Metaverse: Implementation in a Telecollaboration Project for Adult Learner Intercultural Capabilities

Grace Yue Qi

Abstract This chapter proposes a theoretical framework towards humanising language learning in the emerging metaverse. By drawing on language learning as investment (Norton, 2013, 2021) and sociomaterialist theory (Sørensen, 2009), this framework aims to explore learners learning, experiencing, and living their linguistic repertoires in the emerging metaverse where AI and other ubiquitous technologies are situated. The framework centres around "language learners as developing selves" who endeavour to establish, maintain, and strengthen relationships with people, things, and information across time and space in the metaverse era. The proposed framework is implemented in a telecollaboration project involving adult learners of Chinese and English from Aotearoa New Zealand and Taiwan. In considering teaching-as-design, teachers engaged in integrating epistemic, social and (inter/intra)cultural dimensions into their design of the telecollaboration project for learners developing intercultural capabilities. Learners in Aotearoa New Zealand and Taiwan took telecollaboration as an opportunity to explore their inner and outer worlds of being and becoming—navigating and negotiating ways in which language socialisation is not only for meaning making through communication but also for learners to experience a life-course of humanity in the metaverse era.

Keywords Humanising language learning · Telecollaboration · Adult language learners · Intercultural capability · Teaching-as-design

G. Y. Qi (✉)
Massey University, Auckland, New Zealand
e-mail: g.qi@massey.ac.nz

Y. Lan et al. (eds.), *AI-Mediated Language Education in the Metaverse Era*, Chinese Language Learning Sciences, https://doi.org/10.1007/978-981-95-0245-5_3

3.1 Introduction

Language is not only learnt and acquired, but also embodied and lived (Barkhuizen & Strauss, 2020; Ros i Solé, 2016). This is particularly the case today with the emergence of the metaverse (Sparkes, 2021)—an Internet evolution that is significantly transforming the way in which languages are used for interaction and communication across physical and virtual worlds in everyday life. The concept of emerging metaverse has no longer strictly associated language learning with certain physical places, textbooks, teaching materials, or targeted cultures and communities, as language learners today are taking advantage of ubiquitous digital technologies to connect with the world in new ways (Norton, 2013). They have become creative agents who interact with things, people, and information around them (Barkhuizen & Strauss, 2020; Benson, 2021). Their exposure to multiple languages and cultures is a norm by living in a culturally and linguistically diverse society, despite individual learners often aim to pursue language competence of a target language at a time. As language learners explore their linguistic repertoires, their identity is reshaped, likely in the plural form, and continues to evolve, become more dynamic and complex as they are constantly *on the move* with mobile technologies (Benson, 2021; Klimanova, 2021). In the emerging metaverse world, they are also able to live with multiple (digital) versions of themselves, thanks to mixed-realities and social media. However, traditional accounts of language learners in second language acquisition (SLA) and language pedagogy have often focused on aspects of learner identity construction in educational settings and in the contexts of integrating into target cultures and communities (Kinginger, 2004; Norton, 2000; Ros i Solé, 2004), which may be mediated by digital technologies.

In this chapter, I argue for a new theoretical lens to view language learners as *developing* selves (Gibson, 1979; Ingold, 2007) who experience, embody, and live languages, in this globalised and culturally and linguistically diverse world. I see language learners create new meanings and voices that diversify and contribute to human development in the emerging metaverse. To achieve this, a theoretical framework is developed to account for the richness and complexity of language learner identity (re)formed in the metaverse. This framework is particularly foregrounded by language learning as investment (Norton, 2013, 2021) and sociomaterialist theory (Sørensen, 2009). To implement this framework, I present a telecollaboration project, inspired by Whakawhanaungatanga (establish, maintain, and strengthen relationships), a Māori concept, designed for university adult learners who are learning English as an additional language (EAL) in Taiwan and learning Chinese as an additional language (CAL) in Aotearoa New Zealand for their intercultural capability development. This project is informed by design-based research, involving four stages, to manifest the process of how the framework can be utilised to help researchers and language teachers design research-informed learning activities and raise awareness of dynamic configurations of language learners' lives and identities. It is also hopeful that this framework and the implications of implementing the framework in varying contexts fully acknowledge, value, and celebrate the humanity of

language learning in this emerging metaverse today and imagined futures (Hignett, 2021; Kramsch, 2009; Norton, 2013).

3.2 Towards a Theoretical Framework: Humanising Language Learning to Reconceptualise Language Learner Identities

For a long time, researchers in SLA have predominantly regarded language learner identity as developed in conditions of immersion in the target language culture and educational settings (Block, 2006; Kinginger, 2004; Norton, 2000; Pavlenko, 2001; Ros i Solé, 2004). This research agenda, more recently, has shifted to focus on the importance of learner identity development in social contexts and communities of practice (Gibson, 1979) by looking into bilingual or multilingual speakers at home, school, or in a language class (Ros i Solé, 2016). The change of such dynamic in research is based on those individuals who intend to integrate into multilingual communities, or as a consequence of border-crossing and migration (Block, 2007; Coffey & Street, 2008; Kramsch, 2009).

Grounded by poststructuralist theory, Norton (2000) acknowledges individual identity as how a person understands their relationship to the world, how that relationship is developed across time and space, and how they understand and imagine futures (Darvin & Norton, 2015). In her theoretical framing of 'language learning as investment', Norton (2013, 2021) has argued to emphasise the notion of subjectivity that language constructs a sense of self, and that identity is "multiple, changing, and a site of struggle" (Darvin & Norton, 2017). Drawing on postmodern theory, particularly inspired by Pennycook (Pennycook, 2018a, 2018b), identity is "always a gendered subjectivity, [and] rooted also in class, race, ethnicity, and sexual orientation" (Hutcheon, 1989). Both theoretical lenses acknowledge multiple identities that language learners inevitably develop within dynamic temporal and spatial arrangements. Their body, mind, and social practice are inseparable elements that are brought together and interact with one another through language (Kramsch, 2006).

However, little is known in the literature that has examined how individuals' complex, multifaceted identities are (re)constructing, particularly when they are traversing time as well as physical and virtual spaces afforded by the emergence of metaverse in everyday life. This sense of complexities requires a holistic approach to highlight the individual learner's "life capital", defined by Consoli (2022), that is the "repertoire of one's life experiences" (p. 1402). Despite the notion of life capital encourages language researchers, particularly TESOL researchers, to reconsider research ontology and epistemology, this process, in language learning contexts, informs and highlights the importance of learners' identity formation in making connections with their sense of self, other people, things, and information. It is an acknowledgement and a reflection of the 'social turn' in SLA in which Benson (2019) suggests taking socially oriented approaches to understand and research language

learning that view learners as human beings when recruiting them to be part of research projects. Put simply, centring language learners and their learning and lived experience is the key 'social turn' language researchers should work with to emphasise their dispositions and subjectivity as they navigate their being and becoming.

When thinking about language learning environments nowadays, the mainstream research in the field of computer assisted or technology enhanced language learning (CALL/TELL) has highlighted the affordances of ubiquitous technologies for improving learners' language acquisition and (inter)cultural awareness (Klimanova, 2021). Digital technologies mediate learners' identity construction as they engage with multiliteracy practices in virtual realities (Kelly-Holmes, 2019), including using language gestures (e.g., reading, writing, speaking, listening, signing) and semiotics (e.g., images, sounds, graphics, sings, codes) (Love, 2017), communicative genres (Thorne, 2003), multi-modal interactions (Domingo, 2016), and available plurilingual resources (Blommaert & Rampton, 2015; Li, 2017; Ortega, 2017). In this sense, language learners' agency entwins with nonhuman forces, such as human-invented ubiquitous technologies, as reflected in the notion of pattens of relations suggested by Sørensen (2009) that nonhumans and humans coexist in the universe, which emphasises the "understanding of connections and the dynamic practices that emerge" (Carvalho & Freeman, 2022, p. 5). As learners traverse across real and virtual worlds in the emerging metaverse, their online identity is multifaceted and interwoven with their offline condition and dynamically constructed in relation to digital technologies (Bolander & Locher, 2015). With a greater degree of freedom provided by mobile technologies (Ortega, 1997), language learner identities are constantly evolving and (re)shaping—in physical and virtual spaces and learners' imagined futures (Hignett, 2021; Kramsch, 2009; Norton, 2013). Therefore, it is time to illuminate the humanising notion in this emerging metaverse where life is not exclusively human, and that language learners are creating their ways to experience, embody, and live languages in multiple forms (Gee, 2000; Jewitt, 2006; Jewitt & Kress, 2003; Jewitt et al., 2001; Kalantzis & Cope, 2015; Pegrum et al., 2022; Street, 1995).

Towards the humanising language learning in the emerging metaverse framework to understand learner identity construction across time and space, it underscores, from a holistic viewpoint, that identities of individual learners evolve from the moment they are born and develop throughout the dynamic and multiple configurations of their life affairs. Figure 3.1 presents the proposed framework by drawing on the above discussed multiple theoretical lenses. It highlights three interrelated and interdependent levels across time and space. The first level centres around language learners who live in the environmentally, socially, and relationally connected space which involves other people, nonhuman elements, such as things and information, particularly ubiquitous digital technologies. Classrooms and beyond classrooms (Reinders et al., 2022) refer to broadly defined learning environments with assumptions that language learners likely traverse between them, physically and virtually. The second level presents three relationships that language learners develop over time: with one's self (micro), with others (meso), and with the world (macro) (Braidotti, 2020). The relationships contribute to (re)constructing language learner identities in (digital)

social practice where learning takes place temporally and spatially. The third level manifests five key aspects to explore how language learners deal with the world in their own ways. It comprises becomings ("finding yourself" (Lawler, 2012) and "lines of becoming" (Deleuze & Guattari, 2004)), relationality ["relationship with ourselves, with others, with animals, with machines, and with the planet" (Ros i Solé et al., 2020)], language, agencement (Deleuze & Guattari, 2004) (assemblage), and affect (emotion and feeling) (Ros i Solé et al., 2020). The third level highlights the dynamistic and evolving identity construction over time for an individual language learner who constantly and even unconsciously reflects on their own development through interacting with others, and the world around them. All three levels work as an ecological system for one individual language learner as they explore their linguistic repertoires in the culturally diverse world in the emerging metaverse.

The proposed framework aims to provide theoretical guidance to understand multidimensional nature that constructs language learner identities in the emerging metaverse. It is also envisaged that the framework can be implemented in different language learning contexts and environments. Although the framework supports in exploiting ethnographic approaches to investigate life-course lived experience of individual learners to understand their evolving identity construction, it intends to

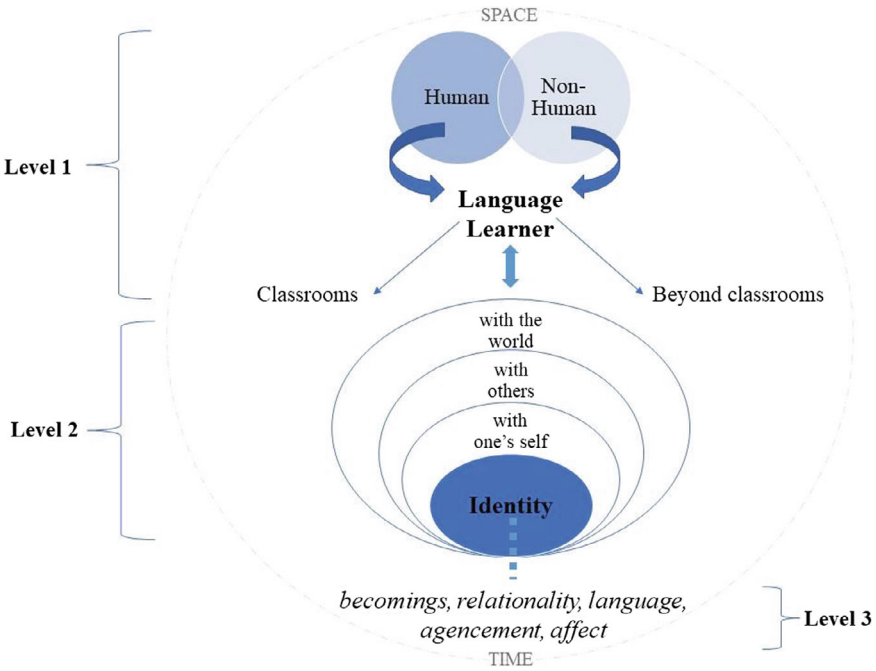


Fig. 3.1 Humanising language learning in the emerging metaverse framework

encourage varying research-informed inquiries in language classrooms in educational settings. To facilitate understanding how this framework might work in practice, a case study presents the implementation of the framework in a telecollaboration project for adult learners of Chinese and English intercultural capability development in Aotearoa New Zealand and Taiwan.

3.3 Implementation of the Framework: A Case Study of Telecollaboration

3.3.1 Context: Identifying Needs

Focusing on intercultural capability development, I implemented the framework in a number of telecollaboration projects with my collaborator in Taiwan in 2022. We have been working on designing different telecollaboration programmes that can be best suited to our language learners of Chinese and English in our respective context. Taking into account the COVID-19 impact, my inquiry into design of a telecollaboration project has been focused primarily on my adult learners' needs and problems in relation to their learning, living, and interacting with the world around them. As most of them are mature aged students, working full time and studying part time, it is essential for them to participate in a telecollaboration project with a sense of meaningfulness, impact, and affect to benefit not only their language and culture learning but also their own personal and professional endeavours. By drawing on Māori concept, I designed a specific telecollaboration project, namely, Whakawhanaungatanga, meaning to establish, maintain, and deepen relationships, for 13 second-semester first year students at the level of intermediate Chinese based in Aotearoa New Zealand who were paired up with 13 English as an additional language (EAL) learners who were enrolled in the Applied English major programme at a central university in Taiwan.

3.3.2 Design: A Project-Based Learning Opportunity

Inspired by Māori colleagues' work (Carlson et al., 2022) in which they focused on improving youth mental health and wellbeing in Aotearoa, framing around Whakawhanaungatanga, their analysis of youths' narratives has allowed them to unpack the concept in detail, which has provided valuable insights into the meaning of each component that has shaped the concept—the Māori way of living, understanding, and knowledge building. Figure 3.2 shows what I reassembled the definitions provided in the article for our telecollaboration project description. Having this briefing of the concept, Whakawhanaungatanga, is believed helpful for participating students, especially those had little knowledge about it in Taiwan. It is also not to

So, what is Whakawhanaungatanga? Well, this could be the first question you find out about yourself and discuss with your partner/s. **In the Learning Portfolio, you are given space to jot down ideas.** Here it is how Māori understand this concept:

> ***AU** refers to 'I', which centres at the heart of individual's world. Their families, relatives and ancestors – **WHANAU** – make (physical, emotional and spiritual dimensional) impact on their life. As **AU** connects with other people, such as friends, colleagues, teachers, mentors, and other community members, **WHANAUNGA** refers to individuals develop and grow further to better understand people, things and events around them. In this sense, individuals develop their own networks – **WHANAUNGATANGA** – which reflect on their initiated and growing relationships with others. In the end, **WHAKAWHANAUNGATANGA**, that is to develop and strengthen relationships as each and every individual lives in the world.*

(Carlson et al., 2022)

Fig. 3.2 Project description: What is Whakawhanaungatanga?

assume that all Aotearoa New Zealand-based students have more knowledge about it. We see each of our participating students equally required such initial support for the Māori knowledge, despite the fact that Aotearoa New Zealand has been advocating for its bicultural and bilingual agenda (Qi, 2021) and moving towards addressing its (super)diversity for multiculturalism and multilingualism (Qi, 2024; Qi et al., 2023).

Drawing on the humanising language learning in the emerging metaverse framework, the design of telecollaboration process was based on students' preference and decision-making. In other words, they were able to negotiate with their partners in choosing topics that can support in developing their Whanaungatanga (relationships) and to decide with possible digital technologies to facilitate their intercultural communication. Most of their interactions were mediated by online meeting platforms of their choice, e.g., Zoom, Google Meet, and social media platforms, e.g., LINE, Facebook Messenger, and Instagram. Their final digital creations were presented in the form of digital storytelling on their chosen digital platforms, e.g., YouTube, Vimeo, to showcase their shared interest and value in pairs through their navigation of understanding the concept of Whakawhanaungatanga.

The project was designed for 13 pairs to work for 9–10 weeks from September when it was the start of the second half of the semester for students in Aotearoa New Zealand and the commencement of a new semester for students in Taiwan. Prior to the execution, they were introduced to the project about three weeks in advance with resources, including Project Description, Grading Criteria, and Learning Portfolio, as well as their agreement to the consent in voluntarily taking part in this research-informed practice. A low-risk ethics approval was received before designing and planning this project in line with the requirement per my university human ethics protocol. Students were asked to provide their preferred contact email address to

Table 3.1 Timeline, steps/tasks, and learning resources of the project

Week	Steps/tasks	Learning resources
1–3	• Provide email address to the Google Sheet • Contact the partner for an initial introduction • Arrange a meeting to discuss the project and procedures to move forward for the next 9–10 weeks	Project Description Grading Criteria Learning Portfolio
4–6	• Focus on the shared interests and values in relation to developing Whanaungatanga (relationships) • Think about and reflect on the process of Whakawhanaungatanga (establish, maintain, and deepen relationships) as collaboration takes place and moves forward • Plan the Digital Storytelling of your shared journey of Whakawhanaungatanga	Learning Portfolio
7–9	• Complete and submit Learning Portfolio individually • Finalise and submit the production/URLs of Digital Storytelling to the course LMS designated area • Submit Telecollaboration Evidence (video recordings, social media interactions, meeting notes, etc.) to the course LMS designated area	Learning Portfolio Digital Storytelling exemplars in the Project Description

share with their project partners on a Google Sheet during the three weeks' preparation time. Initial contacts for each pair were also recommended in this period so they could potentially meet or exchange ideas about how they might proceed their telecollaboration. Table 3.1 demonstrates the timeline of the project suggested to students.

3.3.3 Methods

Aligning with the aim of the project, I employed design-based research approach to undertake this research-informed practice. It involved four stages, including (1) identifying needs (see Context), (2) designing a project-based learning opportunity (see Design), and (3) reflecting (Findings and Discussion). The final stage embraces the data collection throughout the project and data interpreting and reporting, which is discussed in the next few sections. Design-based research, equipped with the three cyclic stages for further project updates, modifications, and redesigns, is an integrated research approach in education (Bannan et al., 2016). It inquiries about theory and learning principles and supports teachers and researchers to create and experiment a new designed solution to tackle complex problems and issues in the learning context (McKenney & Reeves, 2018). Leveraging dynamism, flexibility, and open-ended nature of the practice, design research is situated and supports the interplay of design stages/processes and technology-enhanced language learning phenomenon (Bannan et al., 2010). Applying design research reflects adult language learners' lifeworlds today in which they "have greater control over choices of what, why, where and when they learn" (Sun, 2021, p. 1061).

As a research-informed practice, it is important to collect multiple-sourced data that can support future research to inform improved practice. In this chapter, I focus on the data provided by learners of Chinese based in Aotearoa New Zealand, including their individual completed Learning Portfolio, Digital Storytelling products (pair work), and relevant Telecollaboration Evidence, i.e., video recordings of online synchronous meetings, social media interactions. I undertook the analysis of these qualitative data, mainly thematic analysis (Braun & Clarke, 2019, 2021) for the learning portfolio and multi-modal discourse analysis (Jewitt et al., 2016; Kress & Bezemer, 2023) for the multi-modal data, e.g., Telecollaboration Evidence and Digital Storytelling products. I undertook analyses of all 13 pairs' data, and in the following I present selected findings by drawing on two themes which are also two dimensions in making sense and meaning of Whakawhanaungatanga.

3.4 Findings

The project focuses on two dimensions in relation to Whakawhanaungatanga—one is Whanaungatanga (relationships) to value the importance of having a relationship with the other in the world (Fig. 3.1), and the other is to emphasise the process of establishing, maintaining, and deepening relationships to meet the ultimate goal of Whakawhanaungatanga. The two dimensions are also reflected in the two themes developed from the thematic analysis of Learning Portfolio.

3.4.1 What Does Whanaungatanga Mean to Me?

Learners were asked to research on and exchange ideas in pairs to develop a deeper level of understanding of Whanaungatanga with the initial support of concept understanding in Project Description (see Fig. 3.2). They were encouraged to use visual cues, such as drawings, graphs or images, and texts in reflecting on their understanding of Whanaungatanga as they embarked on the telecollaboration project. One example, as shown in Fig. 3.3, presents a Pākehā (a Māori language term for New Zealanders who have no Polynesian ethnic ancestry, primarily European New Zealanders) student's account of his understanding of Whanaungatanga through an image and further explanation in English texts. While no source is available to the three images, they explicitly illustrate key underlining messages of 'connections' among people (the top two images). The third image is a word cloud that highlights a few keywords, including 'diversity', 'integration', 'respect', 'ethnic', 'people', 'communication', and 'demographic'. Having the two images presented on top along with the word cloud underneath shows an intent to indicate the intersection between them. By reading through written texts by the student, we affirm our assumption that the learner ensures that the choice of images shows the 'signs of learning' to declare "this is what I know" (Kress & Bezemer, 2023, p. 149), which affords the texts to represent

the knowledge the learner has received, interpreted, and reflected to transform and shape his belief about Whanaungatanga. He brings one's identity construction to the light by valuing 'connection with other people and places' and clearly demonstrates his critical thinking about the interplay between individualism and social connections with people and places. As he states, "I think that people who understand their ancestral backstory and history have a clearer understanding of who they are", it provides his level of understanding one's lived experience from the past to the present, and to the imaged future to have a holistic view of one's identity shaping process. All comes down to exhibit the logic of images and texts for meaning and sense making.

A Māori-Chinese student presents and explains Whanaungatanga in his Learning Portfolio (see Fig. 3.4) by using an image to show 'the ties that bind people' and his explanation in multilingual texts. The choice of image is uncertain but can be

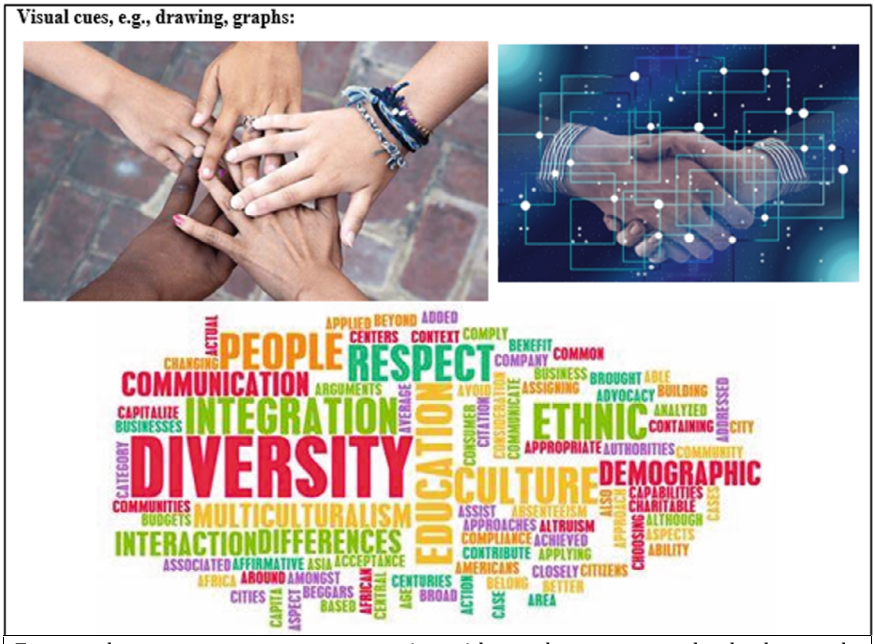
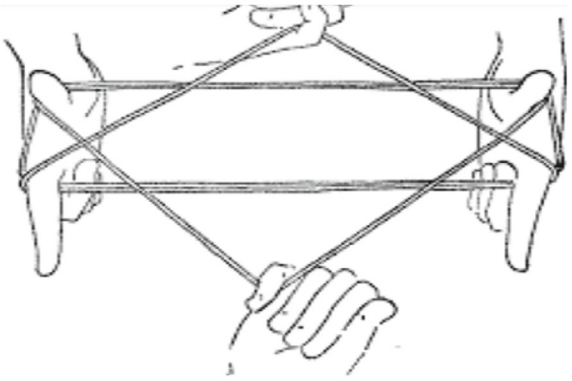


Fig. 3.3 A Pākehā student's multi-modal account of Whanaungatanga

something that resonate with his Māori identity in which he explains in te reo Māori (Māori language), English and Mandarin Chinese to emphasise the importance of 'whānau—family—家', followed by 'whanaunga—relatives—亲属' and to elicit what Whanaungatanga means to him. Differing to the model provided by Carlson et al. (2022) in Project Description (Fig. 3.2), this Māori-Chinese student prioritises whānau and whanaunga over au (self, us), representing a focal point that to him as Māori their immediate connection with others is kinship that refers to family and the wider family. It is explicitly recounted in his final line, "in Māori culture we tend to view all others that we know [through] the 'lens' of the wider family". Born with a hyphenated identity, this student shows his strong attachment and belongingness to his Māori root. Interestingly, as he was paired up with an Indonesian international student studying in Taiwan in telecollaboration, he added 'Gotong Royong' which is the Indonesian concept of 'Whanaungatanga' to remind himself and tell others that the notion of Whanaungatanga exists cross-culturally. However, we can't assume that Whanaungatanga, relationships, 关系 and Gotong Royong represent the shared conceptual meaning within or across cultures. In relation to Whakawhanaungatanga, he elaborates it to make connections between the image and the texts in which he reveals his image choice to present the idea of "those ties that can bind us" for meaning making.



Whanau - Family - 家

Whanaunga - Relatives - 亲属

Whanaungatanga - Relationships - 关系 – Gotong Royong

Whakawhanaungatanga; the process of establishing relationships, relating well to others revealing those Ties that can Bind us

In Maori culture we tend to view all others that we know thru the "lens" of the wider family.

Fig. 3.4 Māori-Chinese student's understanding of Whanaungatanga

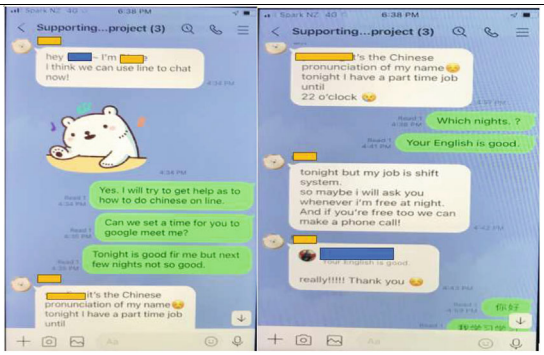
3.4.2 Navigating Whakawhanaungatanga

The analysis of students' different forms of Digital Storytelling products reveals their ways of navigating the concept of Whakawhanaungatanga that might have been distancing from their own cultural beliefs and practices. While knowing what Whanaungatanga is helps to lay a foundation, actioning in Whakawhanaungatanga requires the knowledge gained as well as the collective effort in pursuit for each pair. This process can be challenging and frustrating, as well as rewarding and pleasant.

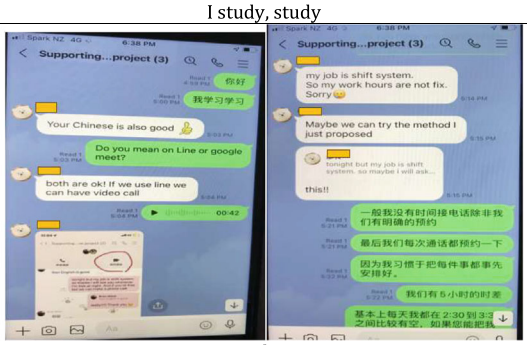
The first example (Table 3.2) shows a site of struggle (Darvin & Norton, 2017) between two students, who have a significant age gap and different ways of living, initiating/establishing their Whanaungatanga. The male student is a Pākehā, mature aged student in his early 80s in Aotearoa New Zealand, while the female student is a young university student in her early 20s from Taiwan. Stating these facts here is to help make sense of why the site of struggle was apparent at this time of their telecollaboration that may have affected their confidence in seeking further progress in maintaining and deepening Whanaungatanga in the telecollaboration experience.

The female student started the social media chat by greeting the male student and introducing herself in both her English and Chinese names. She was quite excited to start this chat with the male student as they finally tackled the technical issues by adding each other and forming this group chat on LINE (a social media platform, commonly used in Taiwan and other east and southeast Asian countries). I knew this because I was requested by the male student as the supporting person being there in case of their needing necessary interventions to facilitate their more instant and effective interactions on LINE. The male student responded but was very task driven by suggesting they set a time for a Google Meet live meeting as soon as possible. While the female student remained polite to tell her partner her Chinese name, she also pointed out that she would not be able to attend the meeting for the night of the day due to her part-time work shifts. She certainly was very aware of a time difference between Taiwan and Aotearoa New Zealand as she implied her work finished by 22.00 Taiwan time which would not work for the male student at all (New Zealand is five hours ahead of Taiwan in daylight savings). However, the male student seemed a little impatient there by following up and querying about "which nights" while was kindly telling the female student "your English is good". The female student prioritised her response on the task as she might have noticed her partner wanted to be clear about their potential synchronous meeting arrangements. And yet, she was more thrilled about his praise to her English skills by using multiple exclamation marks in her reply with a happy smiley face to extend her gratitude towards him, "Really!!!!! Thank you ☺". About 10 min later, the male student switched to Chinese, which was quite a surprise to me, that he probably figured out that he should have used Chinese in the chat—potentially because I was in the group chat observing and he knew he needed to practise Chinese since his partner, the Taiwanese student, had been using English, the language she is learning. His partner also noticed this change and immediately praised his Chinese, "your Chinese is also good (with a thumbs up emoji)". This friendly social cohesion did not prevent the male student

Table 3.2 A site of struggle in embarking on Whakawhanaungatanga



Translation:
Hello
I study, study



Translation:
Hello
I study, study
...
Usually I do not have time to answer phone calls unless we have a confirmed appointment.
Towards the end of each our chat (we need to) schedule our next chat.
Because I am used to organising everything well in advance.
We have a five hours time difference.
Basically everyday I am available between 2.30 and 3.30, if you can...

from continuing his task-driven approach as he returned to the focus to discuss about which platform they shall use for the live discussion. The female student quickly decided on using the LINE videocall function, so they did not have to experience another possible technical struggle in moving between platforms. She instructed where to find the LINE videocall to start a virtual meeting in a screenshot to her partner. However, she followed up to reiterate her point that she could not confirm a meeting time yet because of her flexible work shifts, therefore she suggested "maybe we can try the method I just proposed" and referred to one of her previous messages to the New Zealand partner. This contingent action heightened the site of struggle that the male student's reaction was quite telling. He positioned himself as a sensible, very organised individual, having his preference and lifestyle that he demanded to schedule a meeting like an appointment to be sure of it aligning with his availability.

This positioning of 'self' in negotiation is not uncommon as both have done so in this short social media discourse.

Although their ongoing written modality for communication was LINE mediated, their making connections and establishing Whanaungatanga seems still taking a long way to proceed. In the male student's Learning Portfolio, he reflected on this experience briefly in explaining his understanding of Whanaungatanga and eloquently stated, "我觉得我的 (I think my) relationship with [the project partner] will be limited. 为什么? (why?) 因为(because) age difference, physical difference, that time frame and the forced, artificial nature of the relationship-for-project". To this mature aged student in his life stage when Whanaungatanga must have been meaningful and deeply connected to his real life. The project lasting only about 10 weeks to him was part of the course requirement and unfortunately, not resonating with him to make meaningful connections with his project partner. He preferred real-life, face-to-face connections with people and the world around him. On the one hand, he is clearly not fulfilled by this telecollaboration in terms of navigating Whakawhanaungatanga, but on the other hand, he is well settled in his lifeworlds and knows exactly what he wants and how he can achieve it. Ultimately, he still managed to do his best to work towards the project goal outlined, though, "forced" and "artificial" as he described how he embodied and embraced.

Figure 3.5 presents a digital poster, exhibiting two girls' journey of exploring Whakawhanaungatanga through introducing each of their inner circles of friendship to expanding and making sensible move towards their relationship development over time during telecollaboration. They use 'relationship' as the header of the poster to target the audience of English users. To them, perhaps, Whanaungatanga is relationship as they state, "the state of two or more people being connected" and characterise such 'being connected' refers to "such as family relationship, teacher/student relationship, friendship, romantic relationship". Further, they provide an even broader definition of "friendship" to draw on the wider world to them. They highlight people and ways in which people experience and live "friendship" by building trust, developing compassion, seeking honesty and equality, and becoming independence. Making sense of the world around them also relates to the virtual world where they clearly describe the role of social media plays on the Internet-era new friendship establishment process. However, friendships in real life, such as developed in school and neighbourhoods, do not disappear. They further present their 'shared discussion topics', inquiring about the general sense of friendship, including "what friendship means to us", "how we meet friends", and "what we do with friends" and their reflective accounts of "how we connect us together" mediated by the telecollaboration experience. Photos and the choice of the poster colouring and décor come along with English written texts to showcase their social life in making connections with the others, information around them, and the world. By scanning the QR codes next to their names, the students exhibit further evidence of their interactions during the telecollaboration and how they find ways to focus on friendship to demonstrate their understanding of the complex concept Whakawhanaungatanga. Rather than using the notion of Whakawhanaungatanga, the two students chose to narrow the theme to

friendship that facilitates them to talk about and exchange their own lifeworlds and navigate their connection made throughout the telecollaboration.

3.5 Discussion

The findings of the Aotearoa New Zealand learners of Chinese report on their endeavours with Taiwanese telecollaboration partners in exploring their understanding of two dimensional aspects in relation to the Māori concepts, Whanaungatanga (relationships) and Whakawhanaungatanga (establishing, maintaining, and deepening relationships). Guided by the designed telecollaboration project aiming for intercultural capability development, regardless of individual learners' prior knowledge of Aotearoa New Zealand cultures and histories, they researched on, unpacked, reflected on, and exchanged understandings of Whanaungatanga as they embarked on the journey of Whakawhanaungatanga. The two exemplars of explaining what Whanaungatanga means to learners of Chinese using multi-modal and semiotic resources, e.g., visual cues and images (Jewitt et al., 2001; Kalantzis & Cope, 2015), along with written language(s) as complements, have afforded these learners to effectively communicate their understanding and ways of being (Ros i Solé, 2016). The Māori-Chinese student was particularly aware of his lines of being and becoming (Deleuze & Guattari, 2004; Lawler, 2012) as he already viewed diverse languages and cultures, including his project partner's as the way to make connections and develop Whakawhanaungatanga (Ros i Solé, 2020).

The design of telecollaboration leverages Aotearoa New Zealand-based learners to embody, experience, and live languages, particularly Chinese language as the language they are learning not only in the classroom but also beyond the classroom (Reinders et al., 2022). The seamless connection between the spaces for language learning is facilitated by the ubiquitous technologies they have chosen with their partners to engage over time during the telecollaboration (Level 1 in Fig. 3.1). The pair involving two girls chose to narrow down the concept to 'friendship' which was a strategic move to start their telecollaboration and Whanaungatanga as they were navigating the complexity of Whakawhanaungatanga throughout. Their acknowledgement of virtual and real social engagement afforded them to value their developing Whanaungatanga through this telecollaboration. Little can we know how their identity construction may have been affected by their active engagement over time, however their 'developing self' was evidenced as they embraced the idea of making connections with others, familiar circles, such as friends, or with each other only known from this telecollaboration project, which was significant to their 'life capital' (Consoli, 2022).

By implementing the humanising language learning framework to understand learners' identity construction as they traverse between real and virtual spaces for about ten weeks, the selected pairs reported in this chapter that, while it was a relatively short telecollaboration experience, they have embodied 'developing selves' (Level 2 in Fig. 3.1) as they interact with selves (as language learners), the others

Fig. 3.5 Friendship as a way to navigate Whakawhanaungatanga in a digital poster

(their project partners), and the world (drawing on their own knowledge and ways of being and understanding with open minds for the present and imagined future) (Braidotti, 2020). The multi-modal and semiotic resources have afforded the Pākehā and Māori-Chinese learners of Chinese to socialize with their respective Taiwanese partner to make meaningful connections. While the male Pākehā student who is in his 80's experienced a site of struggle as he was undertaking the telecollaboration project as a part of his long-term investment in learning Chinese at the university, he felt the "forced" and "artificial" Whanaungatanga. Having this affect is understandable as his understanding of Whakawhanaungatanga is based on his knowing and actioning in lived experiences which is inevitably inseparable elements that construct his identity (Kramsch, 2006). There is no need for him to make a real adjustment to his lifestyle and ways of living, thus the configuration of his virtual learning experience and real-life connections with the others and the world has yet to be established. While he grappled with using a new social media and allocating times for synchronous meetings for the project accomplishment goal, he seemed pushing away the 'differences' he has identified to his project partner, including age, lifestyle, and personal interest. He may declare his awareness of intercultural differences and potentially similarities but has no intention of addressing or accommodating them for his enduring experience of learning, living, and embodying through the language.

The nature of the telecollaboration designed as a learning activity in a language programme is a constraint to resonate with the methodological significance of the proposed framework which suggests a holistic study of individual learners' identity construction process. Therefore, little evidence can be drawn on to discuss Level 3 of the framework as I implemented it in this telecollaboration project. Acknowledging the project constraint is part of my reflection to further advancing my endeavours on designing for a project inquiry tailored to humanistic intercultural capability development.

3.6 Conclusion

The humanising language learning framework showcases a new methodological lens to bring out the humanising notion to the long-awaited interdisciplinary research across second language acquisition, language pedagogy, and CALL/TELL. The implementation of the framework in a telecollaboration project illuminates the potential of the framework in expanding to project-based inquiry in language education that language learners are valued as developing selves, making relational awareness, and (re)shaping their identities and engagement with self, others, and the world. As a research-informed practice, it leverages learners and teachers as well as teachers-as-designers to co-construct interculturality integrated into the curriculum. For language teachers and researchers, this chapter also underscores the importance of doing reflexivity to exercise humanness and to (re)position us in the world—as we constantly challenge thinking and understanding and advocate for diversity, equity, and justice vital for humanities (Consoli & Ganassin, 2023a, 2023b). I quote Mario Saraseni's

(University of Portsmouth) work entitled "Language isn't 'alive'—why this metaphor can be misleading", published in the Conversation on 6 June 2023 as follows, in the hope of inspiring future research in this area.

> In reality, languages are integral to what we do as social animals. Societies change through migratory flows, invasions, technological advancement. And as they do so, they also alter the ways in which they use language. Agency rests in the speakers of a language, not in the language itself. … language does nothing. It isn't born, it doesn't grow, it doesn't evolve, it doesn't adapt to the changing environment. It is people who do all these thing (Saraseni, 2023).

References

Bannan, B., Cook, J., & Pachler, N. (2016). Reconceptualizing design research in the age of mobile learning. *Interactive Learning Environments, 24*(5), 938–953. https://doi.org/10.1080/10494820.2015.1018911

Bannan, B., Peters, E., & Martinez, P. (2010). Mobile, inquiry-based learning and geological observation: An exploratory study. *International Journal of Mobile and Blended Learning (IJMBL), 2*(3), 13–29. https://doi.org/10.4018/jmbl.2010070102

Barkhuizen, G., & Strauss, P. (2020). *Communicating identities* (1st ed.). Routledge.

Benson, P. (2019). Ways of seeing: The individual and the social in applied linguistics research methodologies. *Language Teaching, 52*(1), 60–70. https://doi.org/10.1017/S0261444817000234

Benson, P. (2021). *Language learning environments: Spatial perspectives on SLA* (Vol. 147). Multilingual Matters.

Block, D. (2006). *Multilingual Identities in a Global City*. Palgrave Macmillan.

Block, D. (2007). The rise of identity in SLA research, Post Firth and Wagner (1997). *The Modern Language Journal, 91*(s1), 863–876. https://doi.org/10.1111/j.1540-4781.2007.00674.x

Blommaert, J., & Rampton, B. (2015). Language and superdiversity. In K. Arnaut, J. Blommaert, B. Rampton, & M. Spotti (Eds.), *Language and superdiversity* (pp. 29–58). Routledge.

Bolander, B., & Locher, M. A. (2015). Peter is a dumb nut": Status updates and reactions to them as 'acts of positioning' in Facebook. *Pragmatics, 25*(1), 99–122. https://doi.org/10.1075/prag.25.1.05bol

Braidotti, R. (2020). Preface. *Language and Intercultural Communication, 20*(5), 393–396. https://doi.org/10.1080/14708477.2020.1806660

Braun, V., & Clarke, V. (2019). Reflecting on reflexive thematic analysis. *Qualitative Research in Sport, Exercise and Health, 11*(4), 589–597. https://doi.org/10.1080/2159676X.2019.1628806

Braun, V., & Clarke, V. (2021). One size fits all? What counts as quality practice in (reflexive) thematic analysis? *Qualitative Research in Psychology, 18*(3), 328–352. https://doi.org/10.1080/14780887.2020.1769238

Carlson, T., Calder-Dawe, O., & Jensen-Lesatele, V. (2022). 'You can't really define it can you?' Rangatahi perspectives on hauora and wellbeing. *Journal of the Royal Society of New Zealand, 52*(4), 409–425. https://doi.org/10.1080/03036758.2022.2074060

Carvalho, L., & Freeman, C. G. (2022). Materials and places for learning: Experiences of doctoral students in and around university spaces. *Postdigital Science and Education*. https://doi.org/10.1007/s42438-022-00328-x

Coffey, S., & Street, B. (2008). Narrative and Identity in the "Language Learning Project". *The Modern Language Journal, 92*(3), 452–464. http://www.jstor.org/stable/25173070

Consoli, S. (2022). Life capital: An epistemic and methodological lens for TESOL research. *TESOL Quarterly, 56*(4), 1. https://doi.org/10.1002/tesq.3154

Consoli, S., & Ganassin, S. (2023a). Navigating the waters of reflexivity in applied linguistics. In S. Consoli & S. Ganassin (Eds.), *Reflexivity in applied linguistics: Opportunities, challenges, and suggestions* (1st ed., pp. 1–16). Routledge.

Consoli, S., & Ganassin, S. (Eds.). (2023b). *Reflexivity in applied linguistics: Opportunites, challenges, and suggestions*. Routledge.

Darvin, R., & Norton, B. (2015). Identity and a model of investment in applied linguistics. *Annual Review of Applied Linguistics, 35*, 36–56. https://doi.org/10.1017/S0267190514000191

Darvin, R., & Norton, B. (2017). Identity, language learning, and critical pedagogies in digital times. In J. Cenoz, D. Gorter, & S. May (Eds.), *Language awareness and multilingualism* (3rd ed.).

Deleuze, G., & Guattari, F. (2004). *A thousand plateaus.* Continuum.

Domingo, M. (2016). Language and identity research in online environments: A multimodal ethnographic perspective. In S. Preece (Ed.), *The Routledge handbook of language and identity* (pp. 541–557). Routledge.

Gee, J. P. (2000). New people in new worlds: Networks, the new capitalism and schools. In B. Cope & M. Kalantzis (Eds.), *Multiliteracies: Lit learning* (1st ed., pp. 43–68). Routledge.

Gibson, J. J. (1979). *The ecological approach to visual perception.* Houghton Mifflin.

Hignett, M. (2021). *Narratives of Japanese nuclear immigrants in New Zealand: Shinsai experiences, identity, and investment.* (Doctor of Philosophy). University of Auckland.

Hutcheon, L. (1989). *The politics of postmodernism.* Routledge.

Ingold, T. (2007). *Lines: A brief history* (1st ed.). Routledge.

Jewitt, C. (2006). *Technology, literacy and learning.* A multimodal approach: Routledge.

Jewitt, C., Bezemer, J., & O'Halloran, K. (2016). *Introducing multimodality*. Routledge.

Jewitt, C., & Kress, G. (Eds.). (2003). *Multimodal literacy*. Peter Lang.

Jewitt, C., Kress, G., Ogborn, J., & Tsatsarelis, C. (2001). Exploring learning through visual, actional and linguistic communication: The multimodal environment of a science classroom. *Educational Review, 53*(1), 5–18. https://doi.org/10.1080/00131910123753

Kalantzis, M., & Cope, B. (2015). Learning and new media. In D. Scott & E. Hargreaves (Eds.), *The SAGE handbook of learning* (pp. 373–387). Sage.

Kelly-Holmes, H. (2019). Multilingualism and technology: A review of developments in digital communication from monolingualism to idiolingualism. *Annual Review of Applied Linguistics, 39*, 24–39. https://doi.org/10.1017/S0267190519000102

Kinginger, C. (2004). Alice doesn't live here anymore: Foreign language learning and identity reconstruction. In P. Aneta & B. Adrian (Eds.), *Negotiation of Identities in Multilingual Contexts* (pp. 219–242). Multilingual Matters.

Klimanova, L. (2021). The evolution of identity research in CALL: From scripted chatrooms to engaged construction of the digital self. *Language Learning & Technology, 25*(3), 186–204. http://hdl.handle.net/10125/73455

Kramsch, C. (2006). From communicative competence to symbolic competence. *The Modern Language Journal, 90*(2), 249–252. https://doi.org/10.1111/j.1540-4781.2006.00395_3.x

Kramsch, C. (2009). *The multilingual subject*. Oxford University Press.

Kress, G., & Bezemer, J. (2023). Multimodal discourse analysis. In M. Handford & J. P. Gee (Eds.), *The Routledge handbook of discourse analysis* (pp. 139–155). New York Routledge.

Lawler, S. (2012). *Identity: Sociological perspectives* (2nd ed.). Polity Press.

Li, W. (2017). Translanguaging as a practical theory of language. *Applied Linguistics, 39*(2), 261–261. https://doi.org/10.1093/applin/amx044

Love, N. (2017). On languaging and languages. *Language Sciences, 61*, 113–147. https://doi.org/10.1016/j.langsci.2017.04.001

McKenney, S., & Reeves, T. C. (2018). *Conducting educational design research* (2nd ed.). Routledge.

Norton, B. (2000). *Identity and language learning: Gender, ethnicity and educational change.* Pearson Education Limited.

Norton, B. (2013). *Identity and language learning: Extending the conversation.* Multilingual Matters.

Norton, B. (2021). Identity in Language Learning and Teaching. In H. Mohebbi & C. Coombe (Eds.), *Research questions in language education and applied linguistics: A reference guide* (pp. 81–85). Springer International Publishing.

Ortega, L. (1997). Processes and outcomes in networked classroom interaction: Defining the research agenda for L2 computer-assisted classroom discussion. *Language Learning & Technology, 1*(1), 82–93. https://doi.org/10.64152/10125/25005

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

Pavlenko, A. (2001). "In the world of the tradition, I was unimagined": Negotiation of identities in cross-cultural autobiographies. *International Journal of Bilingualism, 5*(3), 317–344. https://doi.org/10.1177/13670069010050030401

Pegrum, M., Hockly, N., & Dudeney, G. (2022). *Digital literacies* (2nd ed.). Routledge.

Pennycook, A. (2018a). *Posthumanist applied linguistics.* Routledge.

Pennycook, A. (2018b). Posthumanist applied linguistics. *Applied Linguistics, 39*(4), 445–461. https://doi.org/10.1093/applin/amw016

Qi, G. Y. (2021). Early Mandarin Chinese learning and Language-in-Education policy and planning in Oceania. In S. Zein & M. R. Coady (Eds.), *Early language learning policy in the 21st century: An international perspective* (pp. 85–107). Springer International Publishing.

Qi, G. Y. (2024). Exploring culturally responsive pedagogy for multilingual education in Aotearoa New Zealand language classrooms: Perspectives of teachers. In L. Gurney & L. Wedikkarage (Eds.), *Language education policies in multilingual settings: Exploring rhetoric and realities in situ* (pp. 107–125). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-57484-9_7

Qi, G. Y., Sun, S. Y. H., & Carvalho, L. (2023). Designing for diversity in Aotearoa / New Zealand Chinese language classrooms. *Language Teaching Research.* https://doi.org/10.1177/13621688231167933

Reinders, H., Lai, C., & Sundqvist, P. (Eds.). (2022). *The Routledge handbook of language learning and teaching beyond the classroom.* Routledge.

Ros i Solé, C. (2004). Autobiographical accounts of L2 identity construction in Chicano literature. *Language and Intercultural Communication, 4*(4), 229–241. https://doi.org/10.1080/14708470408668874

Ros i Solé, C. (2016). *The personal world of the language learner*. Palgrave Macmillan.

Ros i Solé, C. (2020). Lived languages: Ordinary collections and multilingual repertoires. *International Journal of Multilingualism.* https://doi.org/10.1080/14790718.2020.1797047

Ros i Solé, C., Fenoulhet, J., & Quist, G. (2020). Vibrant identities and finding joy in difference. *Language and Intercultural Communication, 20*(5), 397–407. https://doi.org/10.1080/14708477.2020.1784633

Saraseni, M. (2023). Language isn't 'alive' – why this metaphor can be misleading. *The Conversation.* Retrieved from https://theconversation.com/language-isnt-alive-why-this-metaphor-can-be-misleading-205793

Sørensen, E. (2009). *The materiality of learning: Technology and knowledge in educational practice.* Cambridge University Press.

Sparkes, M. (2021). What is a metaverse. *New Scientist, 251*(3348), 18. https://doi.org/10.1016/S0262-4079(21)01450-0

Street, B. V. (1995). *Social literacies: Critical approaches to literacy in development, ethnography and education.* Longman.

Sun, S. Y. H. (2021). Using patterns-based learning design for CALL tasks. *Computer Assisted Language Learning, 34*(8), 1061–1084. https://doi.org/10.1080/09588221.2019.1657902

Thorne, S. L. (2003). Artifacts and cultures-of-use in intercultural communication. *Language Learning & Technology, 7*(2), 38–67.

Grace Yue Qi PhD, FHEA, is a Senior Lecturer in the School of Humanities, Media and Creative Communication, Massey University, Aotearoa New Zealand. Her research interests lie in the epistemological and ontological intersections of language, culture and technology, emphasising the humanistic orientation of language education. Her recent projects and publications include language policy and planning in Australasia, plurilingual education, language teacher agency and identity development, and language learner intercultural communicative competences in real and online multimodal environments.

Chapter 4
AI-Mediated High-Immersion Virtual Reality for Language Learning

Dorothy Chun, Regina Kaplan-Rakowski, Justine Meyr, Uliana Ovsiannikova, Tricia Thrasher, and Ye Yuan

Abstract Based on the proposed theoretical framework of Makransky and Petersen (2021) called CAMIL (Cognitive Affective Model of Immersive Learning) and on an embodied view of cognition for Second Language Acquisition (Atkinson, 2011), this chapter examines the research on how generative artificial intelligence (GAI) can be integrated with high-immersion virtual reality (HiVR) technologies that use head-mounted displays for the purpose of language learning. A review of studies conducted with both commercial and non-commercial HiVR apps revealed that although these immersive environments offer learners new spaces for interaction, most AI-enabled apps do not allow for genuine conversation. Instead, the automatic speech recognition (ASR) capabilities typically only offer users to choose from a set of pre-programmed options, without allowing for spontaneous conversations, leading to a call for expanded AI capabilities (Chun et al., 2022). One promising avenue involves employing GAI chatbots (Huang et al., 2022) that have been found to be helpful for language learning. These so-called "social bots" are designed to be proactive and not just reactive to users (Godwin-Jones, 2023, p. 12). However, the challenge remains as to how to best integrate GAI or large language models (LLM), such as Chat GPT, with HiVR. Generative AI could play the role of a partner, rather than a traditional tutor, and could promote active social learning involving interactions using both the brain and the body (Hwang & Chen, 2023). Furthermore, research and development are needed to take full advantage of the unique affordances of HiVR, specifically

D. Chun (✉) · Y. Yuan
The Gevirtz School, University of California Santa Barbara, Santa Barbara, California, USA
e-mail: dchun@education.ucsb.edu

R. Kaplan-Rakowski
University of North Texas, Denton, TX, USA
e-mail: regina.kaplan-rakowski@unt.edu

J. Meyr
WASC Senior College and University Commission, Alameda, CA, USA

U. Ovsiannikova
MeTabi, Inc., Eugene, OR, USA

T. Thrasher
Immerse, Foothills Ranch, CA, USA

Y. Lan et al. (eds.), *AI-Mediated Language Education in the Metaverse Era*, Chinese Language Learning Sciences, https://doi.org/10.1007/978-981-95-0245-5_4

its fully immersive quality that enhances presence and its multisensory features that allow for holistic embodied participation.

Keywords High-immersion virtual reality (HiVR) · Generative artificial intelligence (GAI) · Embodied learning

4.1 Introduction

This chapter will first review the research on VR for language learning and AI for language learning separately, as the integration of AI and VR is still in its early stages. The few nascent studies combining HiVR and AI will be summarized. Subsequently, underlying theories of VR and AI in education will be presented, revealing a need for refined theories for language learning in particular. Only a few existing HiVR and GAI apps have been used and researched; they will be discussed, along with the effectiveness of their integration of HiVR with GAI. Finally, the chapter will suggest ways in which HiVR and GAI can be integrated to capitalize on the synergies between both technologies to promote authentic interactions and embodied cognition.

4.1.1 Definitions

Virtual reality (VR) has been defined in multiple ways, evolving as new technologies become available. Synthesizing different definitions, Smart et al. (2007) provided a core definition of VR as a system designed to simulate real-life experiences by providing topography, movement, and physics, creating the illusion of "being there." Depending on the different levels of immersion, environments can be described as low-immersion VR (LiVR) or high-immersion VR (HiVR) (Kaplan-Rakowski & Gruber, 2019). LiVR, sometimes referred to as "desktop VR," provides three dimensional (3D) views on computer screens where users navigate in the virtual space through keyboard or mouse manipulations (Kaplan-Rakowski & Gruber, 2019). HiVR intensifies immersion by offering a computer-generated 360° virtual space which users perceive as being spatially realistic (Kaplan-Rakowski & Gruber, 2019). HiVR is accessed by wearing a VR headset (e.g., Meta Quest 3), also known as a head-mounted display (HMD) or goggles, together with the use of the buttons on the headset, a pair of controllers, or haptic systems (Dhimolea et al, 2022). Hand controllers help the user's avatar move around and interact with objects in the VR space. This chapter focuses specifically on HiVR systems and apps.

Artificial intelligence (AI) is a machine-based technique which generates algorithm-based predictions, diagnoses, recommendations, and decisions that influence real or virtual worlds (Hwang et al., 2020). AI covers a broad range of tasks that normally require human intelligence but that can be accomplished by computer

systems and applications. In education, AI has been used as, but not limited to, intelligent tutors, feedback provision, real-time assessment, and progress supervision (Bayne, 2015; Chen et al., 2021). For the purpose of language learning and teaching, a plethora of uses for AI for different aspects of language practice and evaluation have been reported, with some of the most popular topics being automated writing evaluation, intelligent tutoring systems (for reading, writing, and pronunciation), automated error detection, and computer-mediated communication (Huang et al., 2023). In this chapter, we will focus on the particular affordances of natural language processing (NLP) and automated speech recognition (ASR) used by AI chatbots to converse with language learners. Such affordances are integral to VR-AI—assisted simulations, which are defined as "virtual environments that combine virtual reality and artificial intelligence to create immersive and interactive scenarios" (Pitura et al., 2025) and can be particularly useful for language learners to practice conversations in a realistic environment.

4.2 Previous Research

4.2.1 Virtual Reality for Language Learning

Early research on VR compared the effectiveness of LiVR-assisted teaching and traditional teaching methods. Researchers in the 1990s agreed on the affordances of LiVR to enhance classroom engagement, promote intercultural awareness, and lower learners' affective filters (Schwienhorst, 2002). However, computer simulations themselves cannot achieve better learning outcomes without integrated pedagogical instruction. Lee's (1999) study revealed that hybrid simulation, which has expository instructional features, is more effective than pure simulation, which does not integrate any specific directions and explanations. Yet, early studies on VR in education often overlooked the specifics of implementing instructional design within VR platforms, and research on VR-assisted language learning (VRALL) was scarce.

Studies in the first fifteen years of the twenty-first century focused on three major topics: learner interactive communication; learner behaviors, perceptions, and beliefs; and task-based instruction (Lin & Lan, 2015). Research on interactive communication examined learners' metalinguistic awareness and interaction management strategies (e.g., O'Rourke, 2005; Peterson, 2006, 2009). Later studies revealed that non-verbal communication (e.g., kinesics, avatar appearance) was sometimes of more importance to learners than verbal for effective communication in 3D virtual worlds (Wigham & Chanier, 2013). Studies on learner-related topics suggested that LiVR promoted learner autonomy and self-efficacy, reduced anxiety, and fostered creativity (Lin & Lan, 2015). The integration of task-based instruction in LiVR, for example, through built-in activities which guide learners to collaboratively solve problems and complete assigned tasks was shown to have a direct influence on participation and engagement (Lin & Lan, 2015; Peterson, 2012).

Since 2015, the emphasis on HiVR has surged. A systematic review by Dhimolea et al. (2022) revealed major findings of studies from 2015 to 2020 in different language learning domains. Among all the domains, vocabulary learning was most studied (Dhimolea et al., 2022). The results confirmed the affordances of HiVR to promote vocabulary learning by offering contextual exposure (Alfadil, 2020; Garcia et al., 2019; Kaplan-Rakowski & Thrasher, 2025; Legault et al., 2019; Papin & Kaplan-Rakowski, 2022) and interactive exercises (Kaplan-Rakowski & Thrasher, 2024; Nobrega & Rozenfeld, 2019; Xie et al., 2019). Additionally, HiVR was also effective for enhancing speaking (Kaplan-Rakowski & Gruber, 2023a; Pinto et al, 2019; Thrasher, 2022; Xie et al, 2019), listening (Lee, 2019; Ye & Kaplan-Rakowski, 2024), oral communication, that is, engaging speaking and listening skills (Yang et al, 2020), pronunciation (Khatoony, 2019; Thrasher, 2022), and reading (Acar & Cavas, 2020; Kaplan-Rakowski & Gruber, 2023b). The results on L2 writing are hard to generalize due to the different approaches used in VR interventions (Dhimolea et al, 2022). Overall, across different domains, better learning outcomes are possible when learners are given frequent and regular (at least every two weeks) exposure to immersive learning in VR (Dhimolea et al, 2022). To further investigate the implementation and effectiveness of HiVR, future studies should be conducted with larger sample sizes, multiple-time interventions, and triangulated data. For a comprehensive overview of HiVR-assisted language learning research and practice, see Kaplan-Rakowski and Thrasher (2026).

4.2.2 AI For Language Learning

AI in education (AIEd) offers its own unique affordances, as it raises knowledge-gap awareness and allows for personalized and adaptive instruction (Guan et al., 2020). Large-scale reviews have identified trends and topics, both in general and for language education in particular. Reviewing studies on AIEd from 1970 to 2018, Goksel and Bozkurt (2019) identified three major themes: expert systems and intelligent tutoring systems (ITS); AI as an integrated component during instruction; and adaptivity/personalization. The research foci of the first decade of the twenty-first century were on ontology use and knowledge management in ITSs to scaffold problem-solving and example-based learning (Chen et al., 2022). In the later years, the study of learner affect and emotion in diverse scenarios which impact learning outcomes became more prevalent (Chen et al., 2022). The most recent years witnessed a growing interest in AI-assisted personalization aided by AI-integrated techniques such as natural language processing (NLP), learner profiling, ontology, machine learning (ML), and ITS (Chassignol et al., 2018; Guan et al., 2020; Tang et al., 2021; Zawacki-Richter et al., 2019). Increasingly diverse AI technology is extending its application to different educational contexts.

In language education, AI has also gained much attention. AI has unique potential to facilitate language learning as it provides personalized learning systems which allow learners with different proficiency to progress at their own pace (Cha et al.,

2026; Chen et al., 2021). It can also provide location-based contextualized learning environments for real-time authentic interactions (Chen et al., 2019). Hence AI techniques make it possible for self-paced learning anywhere and anytime (Huang et al., 2023). Specifically, AI techniques that are frequently used in language learning are User Modeling, NLP, Natural Language Generation, ASR, and Machine Translation (Gamper & Knapp, 2002). For this chapter, ASR is of particular relevance, especially in the use of chatbots for conducting intelligent conversations (Huang et al., 2023). In the learning of speaking and listening, AI is frequently used to enable learners' verbal interaction with an AI agent (Ayedoum et al., 2019) or a non-player character (NPC) in games (Johnson, 2007), which have great potential in enhancing conversational skills in the target language.

In a systematic review of research from 2015 to 2020 on voice-based intelligent virtual agents in EFL settings, Katsarou et al. (2023) categorized the different types of AI-based virtual agents into four groups. These groups are (1) NLP-based chatbots; (2) embodied conversational agents, which combine verbal and non-verbal communication; (3) intelligent virtual assistants (IVA), also known as intelligent personal assistants (e.g., Amazon's Alexa, Apple's Siri, Google Assistant); and (4) expert virtual assistants, which lack intelligence, except for the pre-programmed information in a particular domain that the program is endowed with. Ten key studies on IVAs for L2 speaking, L2 listening, L2 learners' willingness to communicate, and L2 learners' perceptions of IVAs were presented, with findings of increased L2 oral proficiency (including comprehensible and intelligible communication), decreased anxiety, improved L2 pronunciation, and enhanced listening skills.

In the domain of L2 writing, AI is commonly used via an Automated Writing Evaluation (AWE) system and ITS supported by NLP techniques to diagnose writing errors and promote comprehensive understanding (Huang et al., 2023). In L2 reading, AI is used via ITS to deliver course content and provide learning materials (Huang et al., 2023). AI is also applied to facilitate vocabulary learning by providing contexts of word use (Chen & Li, 2010).

A review specifically of GAI for language teaching and learning discussed the affordances and pedagogical benefits of ChatGPT for English language learning (Kohnke et al., 2023). Intelligent chatbots have the potential to interact with learners in their target language instantly and seemingly realistically; they provide learners with rich input and opportunities to interact as well as modify their output, possibly even to notice gaps in their linguistic production. Especially helpful is the capacity to adjust the linguistic level of input depending on the learner's proficiency level, and to provide a range of vocabulary, expressions, and discussion topics that human partners might lack.

Apart from all the benefits of AI, the literature also reveals challenges faced by AI-assisted language learning. On the one hand, the reliability of AI use in classrooms remains a concern due to different technical issues, such as low accuracy in interpretation and response (Grimes & Warschauer, 2010; Johnson et al., 2017; Roscoe & McNamara, 2013), and disproportionate information generated by biased algorithms (Luan et al., 2020). On the other hand, although teachers and students in general hold positive perceptions toward AI, they may feel reluctant to actually use

AI in language learning due to the uncertainty of its effectiveness and their limited experience with this new technology (Lin et al, 2017; Pokrivcakova, 2019). Last, there are also new concerns regarding privacy and personal information that need to be added to the current code of ethics in AIEd (Zhang & Aslan, 2021).

4.2.3 VR and AI Integration

While the combination of LiVR and AI has been more prominent across various educational contexts including language learning (Ayedoun et al., 2019; Hassani et al., 2016; Mirzaei et al., 2018; Wang et al., 2022; Yang et al., 2022), few studies have looked at integrating AI in HiVR. A combination of VR and AI for language learning can be justified as these technologies have been shown to reduce foreign language anxiety (Katsarou, 2023; Thrasher, 2022) and improve speaking skills in L2 overall (Kaplan-Rakowski & Gruber, 2023a, 2023b; Pinto et al, 2019).

A recent large-scale study on secondary-level students' perceptions of AI agents in a HiVR environment found an overall positive perception of the technology for language learning. Based on questionnaire data, participants were less anxious when conversing with AI agents and they did not experience difficulty comprehending AI agents in a target language (Thrasher et al., 2024b). Another study found that ESL learners enrolled in a secondary school in the U.S. increased their willingness to communicate as a result of practicing speaking English with AI agents in a HiVR environment over a 6-week period (Ovsiannikova, 2024).

At the university level, Ye et al. (2024) created a VR conversation club that allowed international students living in the U.S. to practice English outside of class time. Interview findings revealed that the students' instructors had positive perceptions toward the AI agents, saying that students enjoyed being able to interact in the virtual space with AI agents and that the agents were able to intuitively evaluate what students said during conversations and guide them through speaking tasks.

In another study that examined interactions between AI agents and Chinese learners in an extended reality (XR) system called Cognitive Immersive Language Learning Environment (CILLE), Divekar et al. (2022) found that XR brought visual context to the learners while AI provided opportunities to role-play conversations inside the visual context. Although this is not a HiVR system, it has features that would be desirable for HiVR, namely, that the AI agents can understand multimodal input (they can "see," "hear," and respond in a simultaneous multi-party dialogue), creating more naturalistic conversational experiences; in addition, learners are socially immersed in a realistic environment and can interact multimodally and collaboratively (Divekar et al., 2022).

4.3 Underlying Theories

4.3.1 Theories of VR for Language Learning

Despite the fact that there is a growing body of research on HiVR for educational purposes, surprisingly little work synthesizes the findings into a theoretical framework (Makransky & Petersen, 2021). They thus propose the Cognitive Affective Model of Immersive Learning (CAMIL), suggesting that instructional methods with less immersive media can be generalized to learning in HiVR, but adding the stipulation that media must interact with method. In other words, it is the pedagogical methods that facilitate the affordances of HiVR resulting in increased learning. Specifically, the psychological affordances of *presence* and *agency*, combined with six affective and cognitive factors, lead to HiVR-based learning outcomes (Fig. 4.1).

In second language acquisition, a compelling theory that goes beyond typical cognitivist and sociocultural approaches and that is useful for theorizing about how HiVR can be incorporated into language learning is the *extended, embodied cognition* approach (Atkinson, 2011). This theory, alternately termed the *sociocognitive* approach, posits that "extended cognition conceptualizes mind/brain as inextricably tied to the external environment, while embodied cognition views cognitive activity as grounded in bodily states and action. These two approaches are related because bodies link minds to the world—we experience, understand, and act on the world through our bodies" (p. 599). In addition, Atkinson suggests that "embodied cognition researchers study language processing as a core function of cognition" (p. 605),

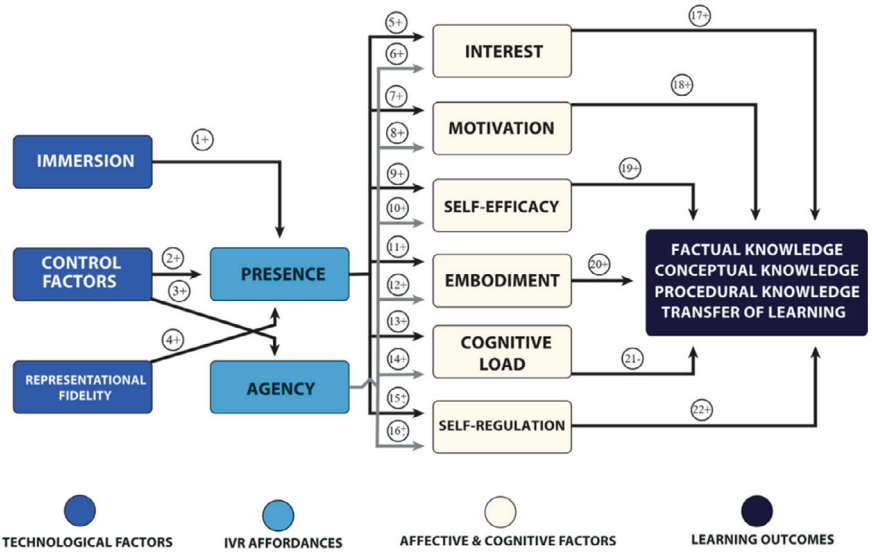


Fig. 4.1 Overview of the CAMIL

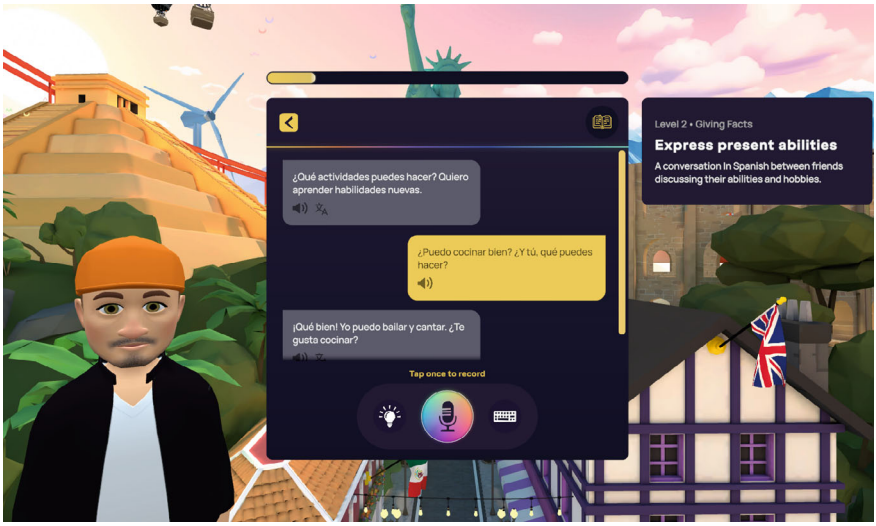


Fig. 4.2 Example of AI conversation in Immerse

making this approach attractive when studying language learners' embodied experiences in HiVR. This view of cognition can be tested in HiVR learning environments that allow users/learners to experience and act in the virtual worlds through their bodies and movements, represented by avatars that are controlled with hand controllers (or keyboards).

4.3.2 Theories of AI for Language Learning

With the rapid rise of using AI for educational purposes in general, the research reviews up to 2025 reveal a lack of studies on AI that are linked to educational theories and thus emphasize the necessity of closely incorporating the application of AI technologies with theory (Chen et al., 2020). In particular, when investigating GAI applications, such as ChatGPT (Chat Generative Pretrained Transformer), researchers have also expressed the need for revisiting pedagogical theories by taking the use of GAI into account and proposing GAI-based educational models or frameworks (Hwang & Chen, 2023). Significant interdisciplinary contemplation and collaboration is needed in this regard.

With regard to AI for language learning, in a systematic review of speech-recognition chatbots for language learning, Jeon and Choe (2023) categorized different types of conversational AI in their Goal-Orientation, Embodiment, and Multimodality (GEM) framework. The framework proposes three main features of chatbots with speech recognition which lead to the emergence of eight types of chatbots for language learning. This framework assesses how the design of conversational

AI, including its specific objectives, virtual embodiment, and communication channels, reinforces affordances of this technology that contribute to language learning. The combination of VR and AI represents a type that combines all three features of the GEM framework enabling such affordances as interactivity, student-centeredness, anxiety-reduction, and communicative authenticity (Jeon & Choe, 2023).

4.4 Currently Available Apps That Combine HiVR and AI

Advancements in AI, particularly in NLP and ASR, have enabled language learners to engage in conversations with AI partners across multiple languages. This capability is available in non-language learning-specific HiVR apps such as *GPT Café*, *Unichat, Virtual Speech,* and *WondaVR.* Moreover, specialized commercial HiVR apps such as *Immerse, ImmerseMe*, and *Mondly VR* claim to incorporate AI within the app to enhance language learning experiences by providing learners with immersive contexts and opportunities for conversational interactions.

4.4.1 Immerse

Immerse (https://immerse.com) is a VR language learning platform that allows users to engage either in live instructor-led group classes, self-paced AI-powered lessons, or informal conversations with AI-avatars in community spaces in Spanish, French, Portuguese, and English. Both live lessons and AI-powered lessons are designed to align with ACTFL levels Novice/Intermediate (low/mid/high) or CEFR levels A1-C1. Research studies conducted on the *Immerse* platform found that learning in *Immerse* improved students' engagement (Lee et al., 2023; Thrasher et al., 2024a), motivation (Kobayashi et al., 2023; Saito, 2023), confidence (Satake & Obari, 2022), anxiety reduction (Saito, 2021; Satake & Obari, 2021), vocabulary learning and retention (Bonner et al., 2023; Kaplan-Rakowski & Thrasher, 2025), and speaking abilities (Dooly et al., 2023; Lee et al., 2023; Saito, 2022). A study of training preservice elementary school teachers to create L2 English lessons in *Immerse* revealed that the teachers benefited from the affordances of HiVR. Their proposed lesson plans capitalized on the sense of *presence* and *agency* that *Immerse* affords, and they were also able to create multiple narratives, social roles, and social interactions in their lessons (Wu & Chun, 2025).

Only three studies have examined how AI chatbots in *Immerse*impact language learning, as AI was only added to its platform in August 2023. In the first study, Thrasher et al. (2024b) explored perceptions of conversing with AI chatbots. Most students (60.9%) believed chatting with AI chatbots helped them learn new phrases and sentences, albeit fewer than half (42.4%) reported that it felt like a real conversation. In the second study, (Ovsiannikova, 2024), 36 refugee high-school students with varying English levels practiced speaking with the AI agents 20-minutes a

week for a 6-week period. The measures of students' willingness to communicate and speaking fluency revealed that students experienced a 19% increase in their willingness to communicate and a 49% increase in the amount they were speaking during AI conversations. Third, Ye et al. (2024) further reported several benefits of using *Immerse*'s AI agents with university-level ESL learners, saying that they led to higher engagement and more enjoyable self-paced learning and personalized practice.

AI chatbots in *Immerse*are powered by ChatGPT-5, or generative AI, which means that learners can converse freely with the bot, as they are not restricted to set responses as on some other platforms (e.g., *Mondly VR*). All chatbots are programmed to complete a specific communicative task with students, such as ordering at a restaurant or negotiating in a business meeting, and students are guided through the activity by an AI agent (Fig. 4.2).

The chatbot can understand learner input that may not be grammatically accurate but is comprehensible. It can also respond to requests from learners (e.g., to slow down and repeat), is programmed to adjust the level of language it produces to the level of the student, and can give the learner feedback as they progress through the activity (Fig. 4.3). Last, *Immerse* includes built-in features to help students while they are conversing with chatbots (e.g., the ability to translate words/phrases from their L2 to L1 and the ability to ask for hints about what to say next in the conversation (Fig. 4.4).

Fig. 4.3 Example of learner feedback in Immerse

Fig. 4.4 Example of hint feature in Immerse

4.4.2 ImmerseMe

ImmerseMe (https://immerseme.co/) is an app currently available for 12 languages that won the 2021 People's Choice Award in Launchpad at the University of Hawai'i. Lessons in *ImmerseMe* are designed to fit ACTFL levels Novice/Intermediate (low/mid/high) or CEFR levels A1, A2, B1, B2. *ImmerseMe* uses 360-degree video technology to surround learners within real-life, culturally authentic locations and scenarios (Fig. 4.5). When wearing VR headsets, learners can travel virtually to many real places around the world and interact with local people. For example, within a few clicks, learners can visit a French café, a German bakery, or a Japanese home for a tea ceremony. Conversations are designed to follow a question-response format; that is, real people ask learners questions (e.g., "What would you like to order?"), then learners reply verbally by selecting one of up to three options provided by the app. If the learners speak comprehensibly enough, the interlocutor in the video will recognize the answer and continue the conversation. Although the scenery and settings are beautifully chosen and are culturally authentic, traditionally, only up to three possible utterances could be recognized by the ASR component, making the conversation limited to the language level chosen by the learner. As of January 2024, *ImmerseMe* offers ChatGPT API integration, allowing for free-flowing, simultaneous

Fig. 4.5 A conversation embedded in a 360-degree (left). A conversation using GAI (right).

conversations. Learners can engage with the chatbot in a similar way to other apps using ChatGPT (Fig. 4.5). This feature is currently (as of 2025) in Beta mode and not yet embedded within HiVR.

4.4.3 WondaVR

WondaVR (https://www.wondavr.com) is a learning and collaboration platform that allows users to edit, share, and access immersive learning experiences based on 2D, 3D, and 360-degree media using VR headsets, mobile devices, tablets, or desktop computers. A recent feature allows for the creation of AI companions or characters who can engage in conversations using GPT4o-mini. The platform has been adapted for language learning at Harvard University, where it has been used to immerse French learners in real-life cultural settings and by students who built their own AI mystery game for their beginning French language course (see Fig. 4.6.)

4.4.4 Mondly VR

Mondly VR is a gamified app that can be accessed via VR headset and purports to be "the first language learning experience with chatbot and speech recognition in the world" (https://www.mondly.com/app). Learners are enticed to "practice real-life conversations in 30 languages." As of 2025, research investigating the effectiveness of MondlyVR is sparse, though studies have shown it to be helpful for EFL listening comprehension (Tai & Chen, 2021) and for vocabulary learning (Chen & Yuan, 2023; Jensen & Cadierno, 2022; Tai et al., 2020). However, in terms of its ASR and AI capabilities, its ability to have meaningful, open-ended conversations is limited. For example, Fig. 4.7 showcases a "conversation" that an L2 learner of Polish had with the *Mondly VR* app. The conversational partner within the app would ask a question or

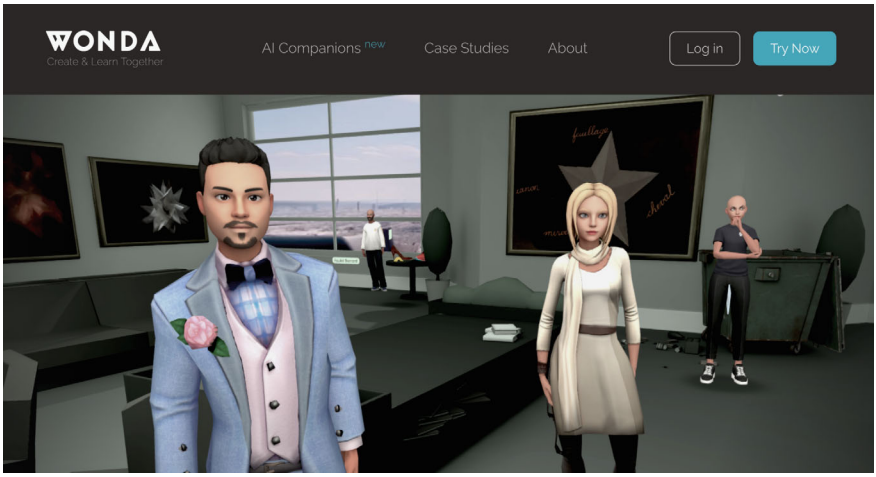


Fig. 4.6 *WondaVR* mystery game to learn French

make a statement, but the learner is limited to selecting one of three pre-programmed responses. These choices are crafted to align with the capabilities of the ASR system, facilitating the understanding of the user's speech. Meanwhile, such communication tends to be restricted and lacks spontaneity.

4.5 Creating Synergies Between HiVR and GAI for Language Learning

As the preceding section signaled, as of mid 2025, both research and development of apps that effectively incorporate both HiVR and AI are in the early stages. This section looks toward the future and suggests avenues for research and implementation of apps that can potentially create synergies between HiVR and AI to realize their full promise and potential (Godwin-Jones, 2023). The challenges and pitfalls of using such technologies together will also be addressed.

4.5.1 Types of Research and Analytical Methodologies Needed

The overarching goal of combining HiVR and GAI is to capitalize on the affordances of both technologies to make immersive learning experiences increasingly holistic and human-like, that is, experiences that mimic authentic human-to-human interaction. Specifically, if the immersive, presence-inducing, embodied learning that is

Fig. 4.7 Snippet of a conversation in *Mondly VR*

facilitated via HiVR can be merged with intelligent AI conversational partners, then language learners will be able not only to interact in their L2 but also have richer sociocultural experiences available at their fingertips.

However, more research is needed to better understand how these chatbots can be designed in a way to most benefit learning. This process includes collecting data based on human-chatbot interactions in HiVR—for example, via video recordings of interactions—and analysis of these recordings from a variety of perspectives. Quantitative studies could measure learning gains for various aspects of language learning, such as competence in speaking, listening, reading, writing, pronunciation, vocabulary, grammar, and syntax. Qualitative studies could focus on pragmatic competence and intercultural communicative competence, learning processes, such as negotiation of meaning, noticing, and self-correction. For non-verbal interactions, written transcripts of the learner-chatbot conversations could be compiled and analyzed accordingly. Triangulation of the results from different data sources is also highly desirable.

Studying the importance of embodied actions specifically for language learning in HiVR environments using one's body and hand controllers will contribute to theorizing about embodied learning (Atkinson, 2011). Possibly connecting gestures to L2 pragmatics or to reinforcing acquisition of specific aspects of L2, such as learning verbs, is another avenue for investigation. There has been preliminary research testing whether interacting with objects in virtual environments contributes to remembering

the names of the objects (Kaplan-Rakowski & Thrasher, 2025), but studies in which learners have used HiVR for an extended period of time, are needed. Similar longitudinal studies which include multiple interventions and/or extended training periods to learn to use GAI, for example, how to write effective prompts in ChatGPT, are essential. Future research on potentially challenging aspects of using headsets in HiVR while interacting with AI chatbots is also necessary. Understanding cognitive load and the distractions caused by wearing headsets and manipulating hand controllers is essential (Mayer et al., 2023).

Research is also needed to better understand what is meant by "intelligent" chatbots. Are they socially aware, and do they interact appropriately, that is, pragmatically and culturally? Do they resemble humans, or are they humanoid intelligent agents or "holographic AIs," and does that have an effect on language learning (Katsarou et al., 2023, p. 79)? Investigating the effectiveness of diverse human-GAI collaborative modes and the multimodal GAI applications has also been suggested (Hwang & Chen, 2023, p. XVI).

To summarize, the types of data analyses that would expand our knowledge and understanding of HiVR combined with GAI for language learning could include:

- Analysis of text transcripts and/or spoken language in video recordings of student-AI-chatbot interactions with regard to:
  - Linguistic accuracy: words, phrases, idioms
  - Negotiation of meaning
  - Speech acts, pragmatic competence
  - Cultural appropriateness
  - Collaborative learning
- Analysis of video recordings of interactions between learner avatars or between learner avatars and AI chatbots with regard to:
  - Gestures, movements, interactions with objects in HiVR that contribute to embodied learning
  - Gestures, movements in HiVR when interacting with AI chatbots that signal/contribute to pragmatic competence and culturally appropriate communication

4.5.2 Challenges in Combining HiVR and AI for Language Learning

As little research has been done on integrating AI in HiVR, there are several challenges associated with using these technologies for learning foreign languages. From the technology design perspective, the biggest challenge is to foster affordances of HiVR and AI simultaneously. On the one hand, HiVR enables embodied and immersive learning, facilitating language learning through gestures and object interaction in an authentic environment. On the other hand, AI affordances of personalization of the learning journey with immediate feedback, contribute to more autonomous

learning. As one way of combining these two technologies, designers may well consider stepping away from a traditional chatbot interface of AI and creating an immersive gamified experience for language learners (Godwin-Jones, 2023).

4.6 Conclusion

The affordances of HiVR include immersion, sense of presence, and embodiment. In the context of language learning, HiVR affordances that play key roles are authenticity, contextualization, learner agency, and safe space (Kaplan-Rakowski, 2024). Many of these affordances align closely with ideal conditions for language learners, mirroring the immersion experience one would encounter in a country where the target language is natively spoken (Vidal & Shively, 2019). Meanwhile, owing to NLP, AI is able to understand, interpret, and generate human language using chatbots, which can be useful for language learners (Shadiev & Liu, 2023) to help them practice speaking, listening, vocabulary, and pronunciation. Using machine-learning algorithms and predictive analytics, AI can also cater for specific user's needs, making their learning personalized (Cha et al., 2026). Personalized and adaptive learning, as well as immediate feedback have been lacking in language instruction. However, as of 2025, we may be experiencing a turning point in which language learners may take advantage of the combined affordances of HiVR and AI. This chapter serves to provide background and a starting point for the forthcoming research on AI in HiVR for language learning.

Although this chapter is primarily about integrating HiVR with specific types of AI, we conclude with reflections on the important connections between research and teaching. In embracing the capabilities of HiVR and GAI technologies, we remain cognizant of the challenges and pitfalls of using such technologies. Such challenges include, for example, the steep learning curves for both teachers and students, the cognitive load factors that can impede rather than enhance learning, access issues of acquiring headsets and having sufficient networking infrastructure, and successfully adapting and integrating the use of these technologies into existing curricula. We acknowledge the necessity of extensive training for instructors and learners alike and of vigilance in ensuring privacy in data collection and analysis. As each educational context is unique, and in some cases, technology (especially when implemented at a large scale) may not be a feasible or helpful option for the learning environment, we urge educators and technology leaders to think critically about the diverse educational contexts they want to enter, and whether integrating technology into these environments is ethical, inclusive, and productive for increasing student learning outcomes.

References

Acar, A., & Cavas, B. (2020). The effect of virtual reality enhanced learning environment on the 7th-grade students' reading and writing skills in English. *Malaysian Online Journal of Educational Sciences, 8*(4), 22–33.

Alfadil, M. (2020). Effectiveness of virtual reality game in foreign language vocabulary acquisition. *Computers & Education, 153*, Article 103893. https://doi.org/10.1016/j.compedu.2020.103893

Atkinson, D. (2011). A sociocognitive approach to second language acquisition: How mind, body, and world work together in learning additional languages. In D. Atkinson (Ed.), *Alternative approaches to second language acquisition* (pp. 143–166). Routledge. https://doi.org/10.4324/9780203830932

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

Bayne, S. (2015). Teacherbot: Interventions in automated teaching. *Teaching in Higher Education, 20*(4), 455–467. https://doi.org/10.1080/13562517.2015.1020783

Bonner, E., Lege, R., & Frazier, E. (2023). Teaching CLIL courses entirely in virtual reality: Educator experiences. *CALICO Journal, 40*(1), 45–67. https://doi.org/10.1558/cj.22676

Cha, E., Ye, Y., Kaplan-Rakowski, R., & Choubey, P. (2026). Personalized adaptive language learning in AI–powered high-immersion virtual reality. In C. Giannikas (Ed.), *Harnessing AI for multigenerational English language learning* (pp. 123–158). IGI Global Scientific Publishing.

Chassignol, M., Khoroshavin, A., Klimova, A., & Bilyatdinova, A. (2018). Artificial intelligence trends in education: A Narrative overview. *Procedia Computer Science, 136*, 16–24. https://doi.org/10.1016/j.procs.2018.08.233

Chen, C. M., & Li, Y. L. (2010). Personalised context-aware ubiquitous learning system for supporting effective English vocabulary learning. *Interactive Learning Environments, 18*(4), 341–364.

Chen, C., & Yuan, Y. (2023). Effectiveness of virtual reality on Chinese as a second language vocabulary learning: Perceptions from international students. *Computer Assisted Language Learning, Ahead-of-Print*. https://doi.org/10.1080/09588221.2023.2192770

Chen, M. P., Wang, L. C., Zou, D., Lin, S. Y., & Xie, H. (2019). Effects of caption and gender on junior high students' EFL learning from iMap-enhanced contextualized learning. *Computers & Education, 140*, Article 103602. https://doi.org/10.1016/j.compedu.2019.103602

Chen, X., Xie, H., Zou, D., & Hwang, G.-J. (2020). Application and theory gaps during the rise of artificial intelligence in education. *Computers and Education. Artificial Intelligence, 1*, Article 100002. https://doi.org/10.1016/j.caeai.2020.100002

Chen, X., Zou, D., Xie, H., & Cheng, G. (2021). Twenty years of personalized language learning. *Educational Technology & Society, 24*(1), 205–222.

Chen, X., Zou, D., Xie, H., Cheng, G., & Liu, C. (2022). Two decades of artificial intelligence in education: Contributors, collaborations, research topics, challenges, and future directions. *Educational Technology & Society, 25*(1), 28–47.

Chun, D. M., Karimi, H., & Sañosa, D. J. (2022). Traveling by headset: Immersive VR for language learning. *CALICO Journal, 39*(2), 129–149. https://doi.org/10.1558/cj.21306

Dhimolea, T. K., Kaplan-Rakowski, R., & Lin, L. (2022). A systematic review of research on high-immersion virtual reality for language learning. *TechTrends, 66*, 810–824. https://doi.org/10.1007/s11528-022-00717-w

Divekar, R. R., Drozdal, J., Chabot, S., Zhou, Y., Su, H., Chen, Y., Zhu, H., Hendler, J. A., & Braasch, J. (2022). Foreign language acquisition via artificial intelligence and extended reality: Design and evaluation. *Computer Assisted Language Learning, 35*(9), 2332–2360. https://doi.org/10.1080/09588221.2021.1879162

Dooly, M., Thrasher, T., & Sadler, R. (2023). "Whoa! Incredible!": Language learning experiences in virtual reality. *RELC Journal, 15*(4), 329–342. https://doi.org/10.1177/00336882231167610

Gamper, J., & Knapp, J. (2002). A Review of intelligent CALL systems. *Computer Assisted Language Learning, 15*(4), 329–342.

Garcia, S., Laesker, D., Caprio, D., Kauer, R., Nguyen, J., & Andujar, M. (2019, July). An immersive virtual reality experience for learning Spanish. In P. Zaphiris, A. Ioannou. (Eds.), *Learning and collaboration technologies. Ubiquitous and virtual environments for learning and collaboration* (pp.151-161). HCII 2019. Lecture Notes in Computer Science, Vol. 11591. Springer. https://doi.org/10.1007/978-3-030-21817-1_12

Godwin-Jones, R. (2023). Emerging spaces for language learning: AI bots, ambient intelligence, and the metaverse. *Language Learning & Technology*, *27*(2), 6–27. https://hdl.handle.net/10125/73501

Goksel, N., & Bozkurt, A. (2019). Artificial Intelligence in Education: Current insights and future perspectives. In S. Sisman-Ugur, & G. Kurubacak (Eds.), *Handbook of Research on Learning in the Age of Transhumanism* (pp. 224–236). IGI Global. https://doi.org/10.4018/978-1-5225-8431- 5.ch014

Grimes, D., & Warschauer, M. (2010). Utility in a fallible tool: A Multi-site case study of automated writing evaluation. *The Journal of Technology, Learning and Assessment*, *8*(6), 1–44. https://ejournals.bc.edu/index.php/jtla/article/view/1625

Guan, C., Mou, J., & Jiang, Z. (2020). Artificial intelligence innovation in education: A Twenty-year data-driven historical analysis. *International Journal of Innovation Studies, 4*(4), 134–147. https://doi.org/10.1016/j.ijis.2020.09.001

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

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

Huang, X., Zou, D., Cheng, G., Chen, X., & Xie, H. (2023). Trends, research issues and applications of artificial intelligence in language education. *Educational Technology & Society, 26*(1), 112–131. https://doi.org/10.30191/ETS.202301_26(1).0009

Hwang, G. J., Xie, H., Wah, B. W., & Gašević, D. (2020). Vision, challenges, roles and research issues of Artificial Intelligence in Education. *Computers and Education: Artificial Intelligence, 1*, Article 100001. https://doi.org/10.1016/j.caeai.2020.100001

Hwang, G.-J., & Chen, N.-S. (2023). Editorial position paper: Exploring the potential of generative artificial intelligence in education: Applications, challenges, and future research directions. *Educational Technology & Society, 26*(2), 1–18. https://doi.org/10.30191/ETS.202304_26(2).0014

Jensen, C., & Cadierno, T. (2022). Differences in mobile-assisted acquisition of receptive and productive vocabulary knowledge: A case study using Mondly. *The Language Learning Journal, Ahead-of-Print*. https://doi.org/10.1080/09571736.2022.2108123

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

Johnson, A., Guerrero, T., Tighe, E., & McNamara, D. (2017). iSTART-ALL: Confronting adult low literacy with intelligent tutoring for reading comprehension. In E. André, R. Baker, X. Hu, M. Rodrigo, B. du Boulay. (Eds.) *Artificial Intelligence in Education* (pp.125–136). AIED 2017. Lecture notes in computer science, vol 10331. Springer. https://doi.org/10.1007/978-3-319-61425-0_11

Johnson, W. L. (2007). Serious use of a serious game for language learning. *Frontiers in Artificial Intelligence and Applications, 158*, 67–74.

Kaplan-Rakowski, R. (2024). High-immersion virtual reality for pragmatics development. In C. A. Chapelle, N. Taguchi, & D. Kadar (Eds.), *The encyclopedia of applied linguistics*(2nd ed.). Pragmatics. Wiley-Blackwell. https://doi.org/10.1002/9781405198431.wbeal1136.pub2

Kaplan-Rakowski, R., & Gruber, A. (2019). Low-immersion versus high-immersion virtual reality: Definitions, classification, and examples with a foreign language focus. In *Proceedings of the 12th International Conference Innovation in Language Learning* (pp. 552–555).

Kaplan-Rakowski, R., & Gruber, A. (2023a). The impact of high-immersion virtual reality on foreign language anxiety when speaking in public. *Smart Learning Environments, 10*(46), 1. https://doi.org/10.1186/s40561-023-00263-9

Kaplan-Rakowski, R., & Gruber, A. (2023b). An experimental study on reading in high-immersion virtual reality. *British Journal of Educational Technology, 55*(2), 541–559. https://doi.org/10.1111/bjet.13392

Kaplan-Rakowski, R., & Thrasher, T. (2025). The impact of high-immersion virtual reality and interactivity on vocabulary learning. *British Journal of Educational Technology*. https://doi.org/10.1111/bjet.13603

Kaplan-Rakowski, R., & Thrasher, T. (2026). Foreign language learning in virtual reality. In J. Plass, G. Makransky, & R. Mayer (Eds.), *Handbook of learning with virtual reality* (pp. xx–xx). MIT Press.

Katsarou, E., Wild, F., Sougari, A., & Chatzipanagiotou, P. (2023). A systematic review of voice-based intelligent virtual agents in EFL education. *International Journal of Emerging Technologies in Learning, 18*(10), 65–85. https://doi.org/10.3991/ijet.v18i10.37723

Khatoony, S. (2019, December). An innovative teaching with serious games through virtual reality assisted language learning. In *2019 International Serious Games Symposium (ISGS)* (pp.100–108). IEEE. Tehran, Iran. https://doi.org/10.1109/ISGS49501.2019.9047018.

Kobayashi, W., Judai, K., Ozaki, T., Kanda, T., Schmidt, M. G., Kamozawa, S., & Tani, S. (2023). The advantages and challenges of VR-enhanced foreign language learning in a Japanese university: A preliminary study. Retrieved from: https://worldcall2023.org/presentation/submission71554/

Kohnke, L., Moorhouse, B. L., & Zou, D. (2023). ChatGPT for Language Teaching and Learning. *RELC Journal, 54*(2), 537–550. https://doi.org/10.1177/00336882231162868

Lee, A. (2019). Using virtual reality to test academic listening proficiency. *Korean Journal of English Language and Linguistics, 19*(4), 688–712.

Lee, J. (1999). Effectiveness of computer-based instructional simulation: A Meta-analysis. *International Journal of Instructional Media, 26*(1), 71–85.

Lee, K., Kwon, O. W., Kim, Y. K., & Lee, Y. (2015). A Hybrid approach for correcting grammatical errors. In F. Helm, L. Bradley, M. Guarda, & S. Thouesny (Eds.), *Critical CALL – Proceedings of the 2015 EUROCALL Conference* (pp. 362-367). Padova, Italy https://doi.org/10.14705/rpnet.2015.000359

Lee, S. M., Yang, Z., & Wu, J. G. (2023). Live, play, and learn: Language learner engagement in the immersive VR environment. *Education and Information Technologies*. https://doi.org/10.1007/s10639-023-12215-4

Legault, J., Zhao, J., Chi, Y. A., Chen, W., Klippel, A., & Li, P. (2019). Immersive virtual reality as an effective tool for second language vocabulary learning. *Languages, 4*(1), 13.

Lin, C. C., Liu, G. Z., & Wang, T. I. (2017). Development and usability test of an e-learning tool for engineering graduates to develop academic writing in English: A Case study. *Educational Technology & Society, 20*(4), 148–161.

Lin, T. J., & Lan, Y. J. (2015). Language learning in virtual reality environments: Past, present, and future. *Journal of Educational Technology & Society, 18*(4), 486–497.

Luan, H., Geczy, P., Lai, H., Gobert, J., Yang, S. J., Ogata, H., Baltes, J., Guerra, R., Li, P., & Tsai, C. C. (2020). Challenges and future directions of big data and artificial intelligence in education. *Frontiers in Psychology, 11*, 1.

Makransky, G., & Petersen, G. B. (2021). The cognitive affective model of immersive learning (CAMIL): A theoretical research-based model of learning in immersive virtual reality. *Educational Psychology Review, 33*(3), 937–958. https://doi.org/10.1007/s10648-020-09586-2

Mayer, R., Makransky, G., & Parong, J. (2023). The promise and pitfalls of learning in immersive virtual reality. *International Journal of Human-Computer Interaction, 39*(11), 2229–2238. https://doi.org/10.1080/10447318.2022.2108563

Mirzaei, M. S., Zhang, Q., van der Struijk, S., & Nishida, T. (2018). Language learning through conversation envisioning in virtual reality: A Sociocultural approach. In P. Taalas, J. Jalkanen, & S. Thouësny (Eds.), *Future-Proof CALL: Language Learning as Exploration and Encounters-Short Papers from EUROCALL* (pp. 207–213). https://doi.org/10.14705/rpnet.2018.26.838

Nobrega, F. A., de Rozenfeld, C. C., & F. (2019). Virtual reality in the teaching of FLE in a Brazilian public school. *Languages, 4*(2), 1–13.

O'Rourke, B. (2005). Form-focused interaction in online tandem learning. *CALICO Journal, 22*(3), 433–466.

Ovsiannikova, U. (2024, May 21-25). *The effect of immersive conversational AI on immigrant students' second language production* [Paper presentation]. *CALICO* Conference, Pittsburgh, PA, USA. https://calico.org/conference-2024/

Papin, K., & Kaplan-Rakowski, R. (2022). A study on vocabulary learning using immersive 360° pictures. *Computer Assisted Language Learning*. https://doi.org/10.1080/09588221.2022.2068613

Peterson, M. (2006). Learner interaction management in an avatar and chat-based virtual world. *Computer Assisted Language Learning, 19*(1), 79–103.

Peterson, M. (2009). Learner interaction in synchronous CMC: A sociocultural perspective. *Computer Assisted Language Learning, 22*(4), 303–321.

Peterson, M. (2012). EFL learner collaborative interaction in Second Life. *ReCALL, 24*(1), 20–39.

Pinto, D., Peixoto, B., Krassmann, A., Melo, M., Cabral, L., & Bessa, M. (2019, April). Virtual reality in education: Learning a foreign language. In Rocha, Á., Adeli, H., Reis, L., Costanzo, S. (Eds.) *New Knowledge in Information Systems and Technologies, 932,* pp. 589–597. WorldCIST'19 2019. Advances in Intelligent Systems and Computing. Switzerland: Springer, Cham. https://doi.org/10.1007/978-3-030-16187-3_57

Pitura, J., Kaplan-Rakowski, R., & Astoska-Wierzba, Y. (2025). VR-AI-assisted simulations for content knowledge application in pre-service EFL teacher training. *TechTrends, 69,* 100–110. https://doi.org/10.1007/s11528-024-024-01022-4

Pokrivcakova, S. (2019). Preparing teachers for the application of AI-powered technologies in foreign language education. *Journal of Language and Cultural Education, 7*(3), 135–153.

Roscoe, R. D., & McNamara, D. S. (2013). Writing Pal: Feasibility of an intelligent writing strategy tutor in the high school classroom. *Journal of Educational Psychology, 105*(4), 1010–1025.

Saito, Y. (2021). Potential and challenges of VR in English education. *KOTESOL Proceedings 2021* (pp. 127–136). https://koreatesol.org/sites/default/files/pdf_publications/KOTESOL.Proceedings.2021.pdf

Saito, Y. (2022). Effects of VR and online public speaking lessons on students' speaking skills. *KOTESOL Proceedings 2022.* https://docs.google.com/document/d/1aqOBAIlFL5m_bPtbDNw4eGY1qAHL5o3kNBQuFzN3kVM/edit

Saito, Y. (2023). Students' creation of VR English lessons: Adopting a constructivist approach. *Japanese Journal of Global Informatics, 3*(1). 31–47. https://researchmap.jp/Yukie_Saito/published_papers/41961966?lang=en

Satake, Y., & Obari, H. (2021). Effects of virtual reality use on Japanese English learners' foreign language anxiety. *ICERI2021 Conference Proceedings* (pp. 1234–1240).

Satake, Y., & Obari, H. (2022). The effects of English conversation lessons in virtual reality on the confidence of Japanese learners of English. *ICERI2022 Conference Proceedings* (pp. 6737–6744).

Schwienhorst, K. (2002). The State of VR: A meta-analysis of virtual reality tools in second language acquisition. *Computer Assisted Language Learning, 15*(3), 221–239. https://doi.org/10.1076/call.15.3.221.8186

Shadiev, R., & Liu, J. (2023). Review of research on applications of speech recognition technology to assist language learning. *ReCALL*, 1:1–15.

Smart, J., Cascio, J., & Paffendof, J. (2007). Metaverse roadmap: Pathway to the 3D web. *Metaverse: A Cross-industry public foresight project.* http://www.metaverseroadmap.org/MetaverseRoadmapOverview.pdf

Tai, T. Y., & Chen, H. H. J. (2021). The impact of immersive virtual reality on EFL learners' listening comprehension. *Journal of Educational Computing Research, 59*(7), 1272–1293.

Tai, T. Y., Chen, H. H. J., & Todd, G. (2020). The impact of a virtual reality app on adolescent EFL learners' vocabulary learning. *Computer Assisted Language Learning*. https://doi.org/10.1080/09588221.2020.1752735

Tang, K. Y., Chang, C. Y., & Hwang, G. J. (2021). Trends in artificial intelligence-supported e-learning: A Systematic review and co-citation network analysis (1998–2019). *Interactive Learning Environments*. https://doi.org/10.1080/10494820.2021.1875001

Thrasher, T. (2022). The impact of virtual reality on L2 French learners' language anxiety and oral comprehensibility: An exploratory study. *CALICO Journal, 39*(2), 219–238. https://doi.org/10.1558/cj.42198

Thrasher, T. (2023). Meeting in the Metaverse: Language learners' insights into the affordances of virtual reality. In D. Cockerham, R. Kaplan-Rakowski, W., Foshay, & M. Spector (Eds.), *Reimagining education: Studies and stories for effective learning practices in an evolving digital society*. Springer. https://doi.org/10.1007/978-3-031-25102-3_17

Thrasher, T., Kaplan-Rakowski, R., Chun, D., Sadler, R. (2024a). Virtual reality: "Awesome", "OK", or "Not so good" for language learning? In B. Bédi, Y. Choubsaz, K. Friðriksdóttir, A. Gimeno-Sanz, S. Björg Vilhjálmsdóttir. & S. Zahova (Eds.), *CALL for all Languages - EUROCALL 2023 Short Papers*. https://doi.org/10.4995/EuroCALL2023.2023.16948

Thrasher, T., Kaplan-Rakowski, R., Ovsiannikova, U., Meyr, J., & Yuan, Y. (2024b). "I can talk to Spanish speakers in Illinois!": Student perspectives on AI-avatar role plays in virtual reality. WORLDCALL 2023 Conference Proceedings.

Vidal, C. P., & Shively, R. L. (2019). L2 pragmatic development in study abroad settings. In *The Routledge handbook of second language acquisition and pragmatics* (pp. 355–371).

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*. https://doi.org/10.1080/09588221.2022.2056203

Wigham, C. R., & Chanier, T. (2013). A Study of verbal and nonverbal communication in Second Life: The ARCHI21 experience. *ReCALL, 25*(1), 63–84.

Wu, Y. J., & Chun, D. M. (2025). An exploration of preservice teachers' use of immersive VR to design English lessons: *Educational Technology & Society, 28*(1), 412–428. https://doi.org/10.30191/ETS.202501_28(1).SP07

Xie, Y., Chen, Y., & Ryder, L. H. (2019). Effects of using mobile-based virtual reality on Chinese L2 students' oral proficiency. *Computer Assisted Language Learning,* 1–21.

Yang, F. C. O., Lo, F. Y. R., Hsieh, J. C., & Wu, W. C. V. (2020). Facilitating communicative ability of EFL learners via high- immersion virtual reality. *Journal of Educational Technology & Society, 23*(1), 30–49.

Yang, H., Kim, H., Lee, J., & Shin, D. (2022). Implementation of an AI chatbot as an English conversation partner in EFL speaking classes. *ReCALL, 34*(3), 327–343. https://doi.org/10.1017/S0958344022000039

Ye, Y., & Kaplan-Rakowski, R. (2024). An exploratory study on practising listening comprehension skills in high-immersion virtual reality. *British Journal of Educational Technology, 55*(4), 1651–1672. https://doi.org/10.1111/bjet.13481

Ye, Y., Thrasher, T., & Kaplan-Rakowski, R. (2024). Establishing a VR club for second language speaking practice. In Names of editors (Eds.), *CALL for Humanity - EUROCALL 2024 Short Papers.*

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education–where are the educators? *International Journal of Educational Technology in Higher Education, 16*(1), 1–27. https://doi.org/10.1186/s41239-019-0171-0

Zhang, K., & Aslan, A. B. (2021). AI technologies for education: Recent research & future directions. *Computers and Education: Artificial Intelligence*. https://doi.org/10.1016/j.caeai.2021.100025

Dorothy Chun is Professor Emerita of Applied Linguistics and Education at the University of California, Santa Barbara and since 2000 the Co-Editor-in-Chief of Language Learning & Technology. Throughout her career, her passion has been researching and exploring how second languages are learned and how technology can be leveraged to enhance the learning of language and culture. Her research areas include L2 phonology and intonation, L2 reading and vocabulary acquisition, computer-assisted language learning (CALL), telecollaboration for intercultural learning, and immersive virtual reality for development of L1 and L2 multiliteracies. She has conducted studies on cognitive processes in learning with multimedia, on whether using visualizations of pitch curves can help L2 speakers of non-tonal languages learn tonal languages and intonation, and on discourse analyses of online intercultural exchanges. In addition to authoring courseware for language and culture acquisition, including apps for Chinese tone pronunciation, she has collaborated on using augmented reality for L2 vocabulary learning and on several virtual reality research projects for language learning.

Dr. Regina Kaplan-Rakowski is an Assistant Professor of Learning Technologies at the University of North Texas (UNT). She is Vice President of Computer Assisted-Language Learning Consortium (CALICO). Dr. Kaplan-Rakowski's key research interests are at the intersection of three areas: immersive learning technologies (e.g., virtual reality, augmented reality, mixed reality), emotional responses to learning technologies (e.g., immersion, performance anxiety, motivation, engagement), and computer-assisted language learning (CALL). Within CALL, Dr. Kaplan-Rakowski studies how learning languages can be facilitated using generative artificial intelligence, high-immersion virtual reality, immersive stereoscopic 3D pictures, 360-degree pictures, and other immersive technologies. She serves on the Editorial Board of Educational Technology & Society (ET&S) and the Journal of Technology and Teacher Education (JTATE).

Justine Meyr is currently the Director of Curriculum Development for Western Association of Schools and Colleges (WSCUC). Her principal research and design initiatives focus on intercultural communication and community building in online, in-person and blended learning environments. She is currently working on the creation and assessment of online training courses for adult learners using adaptive learning platforms. She has published in journals such as the *Journal of Second Language Pronunciation* and the *Portuguese Language Journal* and has served as an editor for the *Journal of Virtual Exchange*. In her free time she enjoys traveling, making pottery, hiking and triathlons.

Uliana Ovsiannikova holds a master's degree in teaching English as a Second Language (TESL) with a minor in Game Design from the University of Illinois Urbana-Champaign. During her time as a Research Assistant on a Meta-funded project, she explored the intersection of Virtual Reality and AI to enhance language education. Her academic work focuses particularly on supporting immigrant youth, fostering their language production and willingness to communicate in a foreign language through innovative educational technologies. In her current position as a Learning Solutions and Engagement Manager at MeTabi, Inc., Uliana continues to explore affordances of immersive conversational AI for language learning in various academic settings. As a CALICO Graduate SIG Co-Chair, Uliana is also deeply involved in the academic community, organizing workshops and contributing to discussions on technology's role in language education.

Tricia Thrasher (Ph.D.) obtained her Ph.D. from the department of French and Italian at the University of Illinois at Urbana-Champaign in May 2022. She is currently the Research Manager at Immerse, a VR language learning company. Her research interests include immersive technologies for language learning, foreign language anxiety and oral production, and subjective and objective anxiety measurements. She has published in journals such as the CALICO Journal and Foreign Language Annals and is currently the Chair of the CALICO Immersive Realities Special Interest Group.

Ye Yuan is a Ph.D. candidate with interdisciplinary emphases in Applied Linguistics and Language Education at the Gevirtz School of Education at the University of California Santa Barbara. Ye earned her B.A. in TCSOL from Beijing Language and Culture University in China, and her M.A. in TESOL from the University of Southern California. Currently, her research interests lie in bilingualism, heritage language maintenance, and the use of technology in language learning.

Chapter 5
A Comparison of Voice-Based Versus Text-Based Communication with Non-character Players During Complex Tasks on the 3D Virtual World Platform of Second Life: Preliminary Student Feedback and Design Implications

Scott Grant

Abstract The ongoing advancement of computer hardware and software technology over recent decades has attracted a lot of interest from foreign language educators and researchers. As new and improved technologies become available, educators are looking at a multitude of ways to leverage these technologies to enhance language learning within the limitations of the traditional classroom environment. Among the technologies of interest, virtual worlds, AI, and speech-to-text (STT) or automatic speech recognition (ASR) are three that have the potential to be combined to provide learners with situated learning in immersive simulations of real-life scenarios. This chapter reports on a pilot study involving 6 learners of Chinese as a foreign language (CFL) who while engaged in two lessons in the 3D virtual township of Chinese Island in Second Life interacted with non-player characters (NPCs) in both text-based chat and voice chat. The purpose of the study was to illicit feedback from participants on perceived or actual differences, challenges, benefits, or disadvantages of both forms of communication. Analysis of this feedback is intended to provide the basis for making necessary adjustments to lesson design and technology implementation for a future larger scale study comparing the relative efficacy of voice-based and text-based interaction with the Chinese Island NPCs. This chapter first gives some background on the technologies used and on the larger teaching and learning context in which they have been combined and utilised. Also discussed are the application of situated learning and simulation in a 3D virtual environment, as well as the key differences between chatbots and NPCs. Finally, some of the key findings from the pilot study are presented, which will serve as a reference for future study designs and pedagogical practice and design.

S. Grant (✉)
Chinese Studies, Faculty of Arts, Monash University, Clayton, Melbourne, Australia
e-mail: scott.grant@monash.edu

Y. Lan et al. (eds.), *AI-Mediated Language Education in the Metaverse Era*, Chinese Language Learning Sciences, https://doi.org/10.1007/978-981-95-0245-5_5

Keywords 3D multiuser virtual environments · Non-player characters (NPCs) · Automatic speech recognition (ASR) · Situated learning

5.1 Introduction

5.1.1 From "Metaverse" to Generative AI

3D virtual world and virtual reality platforms have been around for many years now, and a significant body of research has developed (and is growing) around the use of these platforms for educational purposes. However, during the period when the COVID pandemic had a profound impact on the structure of work and education, i.e. 2020 to 2022 (CDC Museum COVID-19 Timeline; Reimers, 2022) the concept of the "metaverse" heavily promoted by Meta (formerly Facebook) (Milmo, 2021) began to dominate and subsume public awareness of these other well-established platforms. Along with the hype predicting that we, the general public, would be spending significant amounts of time living, working, studying, socialising, and playing in the "metaverse" in the future (Bojic, 2022; Dundas, 2021; Dwivedi et al., 2022) also came a lot of criticism. From 2022 onwards reports began appearing in the media about the failure of Meta's vision of the metaverse (Hunt, 2023; Manavis, 2023), Meta's laying off technical staff (Vanian & Levy, 2023; Vanian & Goswami, 2023), and of the demise of the "metaverse" in general (Mac Ghlionn, 2023). Social virtual worlds such as Second Life and VRChat and game worlds such as Roblox and Fortnite, however, continued to thrive.

By the end of 2022, there was a new "kid on the block" that stole the limelight from the metaverse—generative AI. After a number of years of development (GPT-1 in 2018, GPT-2 in 2019, GPT-3 in 2020) (Marr, 2023), the developer of ChatGPT released a demo of its powerful generative AI model in November 2022. In all the excitement about this new, powerful AI, a multitude of possibilities about its uses began to be canvased, from education (Alphonso, 2023; Rudra, 2023) to retail (Standish, 2023), to cybersecurity (Taulli, 2023). At the same time, there was a great deal of uncertainty, even fear, of what problems such a powerful AI could cause. Warnings abounded of the extinction of humanity (Collier, 2023), threats to the health of millions (Gregory & Hern, 2023), cheating and ethical concerns in the field of education (Davis, 2023a, 2023b), job losses for millions (Valance, 2023), along with reports of schools and universities around the world banning the use of generative AI in response to these concerns (Davis, 2023a, 2023b; De Clercq, 2023).

5.1.2 A Brief History of Chatbots and Generative AI

While this powerful new generative AI technology stole attention from the nascent virtual dimension of human activity called "the metaverse" and was being rolled out by an increasing number of companies (Google Bard, Microsoft Bing (now Copilot), etc.), AI is not new. It has a long history of exploration and development stretching back to the beginning of the twentieth century (Toloka.ai, 2023). In terms of general AI, mathematician Alan Turing, based on his own exploration of AI, proposed his Turing Test to measure the intelligence of machines (computers) in the early 1950s (Hawley, 2023), a test that is still considered by some to be an important yardstick today (Somoye, 2023). Neural networks, the foundation for today's generative AI, began being explored in 1940s, with McCulloch and Pitts developing a conceptual model of an artificial neuron, a fundamental unit of deep neural networks (Chandra, 2018) that was later built on by Frank Rosenblatt's classical perceptron proposed in 1958 and further refined by Minsky and Paper in 1969 (Chandra, 2018). The birth of AI as a field of study is attributed to the Dartmouth Summer Research Project on Artificial Intelligence held in 1956 in the USA which brought together researchers from a wide range of disciplines to discuss the possibility of thinking machines (Hawley, 2023).

Among the first primitive working examples of a Natural Language Processing (NLP) AI was a program called ELIZA, a chatbot created by German American scientist Joseph Weizenbaum in 1966 (Tarnoff, 2023). The program was based on simple keyword pattern matching and set generic responses, which enabled the chatbot to communicate with humans in natural language (Toloka, 2023). This type of pattern-matching AI was further developed by Dr Richard Wallace when he created the Artificial Intelligence Markup Language (AIML) used to create the A.L.I.C.E (Artificial Linguistic Internet Computer Entity) chatbot (Pandey & Sharma, 2023). A.L.I.C.E has won the annual Loebner Prize Competition in Artificial Intelligence three times. The Loebner Prize is an annual competition based on the Turing Test (Powers, 1998). The aim of the competition is to see if competing chatbots can convince a panel of judges that they are communicating with a human interlocutor. The Loebner Prize has, however, attracted a lot of criticism over the last few decades due to the highly constrained circumstances under which they are conducted, conditions that Turing did not include in his original design of the test (e.g. time limits on communication with the AI) (Powers, 1998) and the competition is largely ignored by the AI community (Shieber, 2011) although the Turing Test is still seen as one yardstick by which to measure the performance of generative AI like ChatGPT (James, 2023).

Research into machine learning, a subfield of AI, has been around since the mid-twentieth century. The heart of machine learning is the use of algorithms by computers to learn and make decisions from data supplied to them (Khan, 2023). One of the more profound breakthroughs in machine learning came about through the resurgence of neural network research and development in the twenty-first century leading to the advancement of deep learning which allows the training of neural networks in multiple layers (Khan, 2023). The hallmark of this renewed focus on neural networks

occurred in 2012 when a deep convolutional neural network called AlexNet won the ImageNet Large Scale Visual Recognition Challenge (Khan, 2023) with image recognition rates much higher than other competitor AI. The leap forward in deep learning brought about by AlexNet's architecture is sometimes credited with bringing deep learning to adjacent fields like natural language processing (Wei, 2019). Prior to the advent of generative AI (particularly before the release of ChatGPT in late 2022), the majority of chatbots in, for example, the therapeutic area (Pandey & Sharma, 2023) were based on the pattern matching and set response approach together with set dialogue management frameworks (often using AIML) (Dsouza et al., 2019). This approach is also known as the retrieval-based approach. The most significant difference between retrieval-based chatbots (like A.L.I.C.E) and generative-based chatbots (like ChatGPT) is that the former are restricted to responding to inputs that best match the outputs that they already know. Generative AI-based chatbots in contrast have the ability to generate new text in response to inputs (Pandey & Sharma, 2023).

5.1.3 Natural Language Processing (NLP) and Automatic Speech Recognition (ASR)

Natural Language Processing, or NLP, is a subfield of AI focused on interaction between humans and computers through the use of natural human language. It has a long history dating back to the late 1940s, early 1950s (Jones, 1994). The aim is to enable computers to understand and respond to human language in a meaningful and contextually relevant way (Simon Says, 2023). Human–computer interaction (HCI), the study of how humans interact with computers, is a multi-disciplinary field of research with a history stretching back nearly 80 years. While a separate field of research and development endeavour, NLP research has contributed significantly to the field of HCI in terms of making it possible for humans to communicate with machines in ways that emulate human–human interaction (Manaris, 1998). In terms of making interaction with computers more "human-like" Automatic Speech Recognition (ASR), used by a wide range of modern applications, has made a significant contribution. NLP is key component of speech recognition technology in that after audio information is processed NLP techniques are used to decode the text produced into natural language and to recognise the context in which they are used (Tyagi and Bhushan, 2023).

What is speech recognition (SR)? A simple explanation is that our voices are turned into written words through a transformation process (Reid, 2023). A synonymous term, Automatic Speech Recognition (ASR), indicates that speech is automatically machine processed. ASR is able to support continuous speech, can work independently of the speaker, and in some cases has achieved human levels of performance (Kodish-Wachs et al., 2018). ASR has the potential to greatly speed up ongoing

research into NLP and the development of practical tools in a range of fields including the clinical diagnosis of schizophrenia (Ciampelli et al., 2023).

The first machine capable of recognising human speech was called Audrey and was invented by Bell Labs in the USA in 1952. Audrey was by modern standards extremely large, practically taking up half a room of space, yet it could only recognise 9 words (the numbers 1–9). Ten years later this number had increased to 16 words with the creation of a smaller machine by IBM. Both these systems split spoken words into phonemes, the smallest unit of sound to have any meaning which they then reassembled the phonemes as complete words. A major leap forward in terms of accuracy and vocabulary size came in the 1970s and 80 s with speech recognition systems moving towards the use of probabilistic guessing of what words were most likely to follow each other using n-gram language models and Hidden Markov Models (Dictate.it, 2021). The advent of deep learning in the early twenty-first century when applied to speech recognition saw significant improvements in speed and accuracy over the previous approaches and Deep Neural Networks (DNNs) are now commonly used in devices such as Siri, Alexa, Cortana, and so on (Dictate.it, 2021) and programs like Microsoft Automatic Speech Recognition (Li, 2014) and Apple Dictation, all of which use ASR.

5.1.4 3D Multiuser Virtual Environments

While the concept of a single "metaverse" has lost traction in the popular imagination, multiuser social, gaming, and educational 3D virtual environments still attract a large and growing user base. On the gaming front, Roblox, a multiuser 3D virtual environment platform first launched in 2006 that allows users to play dozens of user-created games or create their own games (Garrett, 2023), had more than 65 million daily active users in September 2023. Initially a platform designed for children, developers are now making in-game experiences focused specifically on adults including meeting and dating. In addition, the platform now includes generative AI tools to aid developers in creating new experiences (Peters, 2023). Fortnite, another multiuser 3D open-world game platform, as of February 2023 had 250 million average monthly players (Ibrahim, 2023). While the core play mechanics of Fortnite is a battle royal with 100 players at a time dropped on to a virtual island to fight it out to be the last one standing, the game has distinguished itself in recent years by holding a number of large-scale live concerts by major real-life popular performers (Maddock, 2023). Following in the footsteps of other multiuser gaming platforms, Fortnite also provides users with a creation-focused toolset that allows them to create custom game modes and rulesets with up to 16 other collaborators (Zamora, 2023).

Social platforms like Second Life and VRChat while having smaller user bases than the aforementioned gaming platforms, still attract significant numbers of regular users. Launched in 2003 by Linden Labs, Second Life allows users to interact with each other via avatars within a virtual world (Dwivedi et al., 2022) and to create whatever their heart desires. The virtual world space has been used by players to

build mansions, to stage art exhibitions and theatrical performances, to run virtual stores, and even to conduct educational activities. User interaction occurs via text chat, voice chat, and body language. At the height of its popularity in the mid-2000s, Second Life attracted around a million monthly users (Parkin, 2023). Current figures indicate there are around 600,000 monthly active users (Wagner, 2023a). With its own virtual currency Linden Dollars which are fully convertible to US dollars, Second Life boasts an annual GDP of US$650 million in sales of virtual items created and purchased by users (Parkin, 2023). VRChat, launched in 2014, is a virtual reality (VR) platform that can be accessed via computer screen and VR headsets. Users interact with the virtual environment and other users via avatars locked in first person view. The use of VR headsets gives users an even stronger sense of presence within the virtual environment and augments their sense of ownership over their avatars (Deighan et al., 2023). It is considered one of the most popular VR platforms with numbers during the COVID pandemic period reaching 22,000 daily users and a total of over 4 million users (Deighan et al., 2023). As with Second Life, most of the content in VRChat is user generated, and there are approximately 50,000 user-created worlds on the platform (Deighan et al., 2023). Interaction between users occurs mainly through voice chat (text chat is still very limited in VRChat), although there is a significant community of users who through preference or disability use sign language to communicate (Wagner, 2023b). In addition to providing a virtual space for social connection, researchers have found VR applications like VRChat provide users with opportunities to actively support their mental health and wellbeing in daily life (Deighan et al., 2023).

5.2 Literature Review

5.2.1 Text-Based Chat

Text-based chat in synchronous computer mediated communication (SCMC) has been referred to in the past as a "conversation in slow motion" (Beauvois, 1997). Unlike voice-based conversations where interaction tends to be rapid, utterances disappear immediately upon being emitted, and long response times potentially create a feeling of awkwardness, in text-based conversations participants can take time to read and absorb what the other party is saying and formulate their own response without creating a feeling of awkwardness or disjointedness. This is due in part to utterances lingering on the screen or being easily recalled via "chat history," which has been called "persistent conversations" (Thorne, 2008). This is particularly useful for introductory level foreign language learners who have limited vocabularies, little experience of using the foreign language communicatively and may find the rapid pace of verbal communication overwhelming. Further, Beauvois posits this kind of communication as being a "hybrid phenomenon that falls somewhere between writing and speaking" due to the relatively "rapid nature of the electronic exchanges"

(1998, p. 198) (in contrast with asynchronous communication). Text-based SCMC has the additional benefit of being easily, and in some cases automatically, logged, providing rich data for learner reflection and teacher/researcher/developer analysis post-conversation (Thorne, 2008).

A number of studies have shown CMC, particularly text-based conversational interaction, as being an effective method of increasing learner production (Sands, 2006; Sykes, 2005) and facilitating noticing (Schmidt, 1990) (of gaps, problems, or a need for revision of messages posted by participants) through the persistent nature of text-based communication (Erickson, 1999; Thorne, 2000, 2008), both of which are seen as beneficial for language acquisition.

5.2.2 ASR

In their meta study of over 350 studies covering a range of technologies used in foreign language learning and teaching across recent decades, Golonka et al. found that support for the claim that technology made a measurable impact on foreign language learning came from studies that in particular looked at automatic speech recognition (2014). They noted that while error detection reliability is an issue, students' attitudes to using ASR-based programs was positive (Sun, 2023) and that the inclusion of chat in foreign language learning increases language production and its complexity (Golonka, 2014). It should be noted that most of the ASR programs referred to in the studies included an error detection and pronunciation correction component in addition to speech recognition, functionality that was not part of the pilot project described in this chapter. Feedback was a significant issue raised by the participants in the pilot project and potential solutions for a larger scale study are discussed at the end of the chapter.

5.2.3 Foreign Language Pedagogy

5.2.3.1 Foreign Language Versus Second Language Classroom

The pilot project discussed in this chapter was conducted at a major Australian university that has one of the largest Chinese studies, language, and culture departments in Australia. As such, all of the language teaching that occurs during the normal semester occurs under the constraints of a foreign language classroom environment and a tertiary curriculum structure. Unlike second language classrooms that are embedded in an environment in which the mainstream language is the language being learned and opportunities for contextualised situated learning and practice abound (Falvey, 1988; Gass & Selinker, 2008), the mainstream language spoken outside the foreign language classroom is not the one being learned and as such opportunities have traditionally been much more limited. In their meta-analysis of 97 studies published in

The Modern Language Journal between 2001 and 2014, Collins and Muñoz found that the differences between the two learning contexts include "target language use outside the classroom and the quality, quantity and diversity of input" (2016, p. 134).

5.2.3.2 Limitations of Formal Tertiary Curriculum

Adding to the problem is the restrictive nature of many conventional tertiary curriculum formats. The shape of this format is largely pre-determined by the institution and leaves little flexibility in terms of course structure, number of contact hours, student/teacher ratios, forms of assessment, etc. A typical structure might consist of a large group lecture (either live or recorded), follow up small group tutorials, language labs, and maybe some regular online assignments/exercises/assessment via an LMS. Assessment often takes the shape of regular quizzes, larger mid-semester tests, and end-of-semester high-stakes exams. It is important, but not ideal or sufficient for, multiliteracy, multicultural, or foreign language learning.

5.2.3.3 Contextualised Situated Learning and Pragmatic Competence

Contextualised situated learning has been shown to be an effective means of achieving pragmatic competence. Taguchi (2018) says that pragmatic competence can be broadly defined as "the ability to perform language functions in social interactions" and draws heavily on context. Further, Taguchi argues that there is "a general consensus that the classroom lacks opportunities for pragmatic learning because the classroom discourse does not provide representations of a variety of communication situations" (2015, p. 8) Learning in authentic contexts facilitates the generation of meaningful language through the provision of rich and multimodal information and helping students connect newly acquired knowledge with existing knowledge. Studying in a country where the language being learned is widely spoken, for example, can provide a contextualised and situated learning environment. This also presents many challenges, including cost, time restrictions, family commitments, and can actually be demotivating for some students (Allen & Herron, 2003). One alternative is creating an interactive multiuser 3D virtual learning environment that can provide authentic learning context and associated tasks for situated learning. Taguchi argues that virtual environments are a hybrid learning environment that makes up for the shortcomings of both the formal classroom and study abroad. Further, virtual environments combine specific elements of pragmatic learning that include context, functional language use and interaction, experiential learning, and authenticity (Taguchi, 2015).

5.2.4 Context and Background of the Current Study: A 3D Virtual World Platform in Second Life

The 3D multiuser virtual environment used in the pilot study discussed in this chapter is a virtual township on a virtual island in Second Life that was established in 2008. Second Life offers a fully customisable virtual environment that can be tailored to specific educational and pedagogical needs. While there are many platforms and apps available that utilise 3D single and 3D multiuser virtual environments (3DMUVE) to provide language education, they often don't allow much/any customisation and have a set pedagogical approach that may or may not suit the specific teaching situation.

In an attempt to design lessons that provide opportunities for authentic communication within the confines of a foreign language classroom, an approach known as task-based language learning (TBLL) was adopted. TBLL involves communication activities using language that is meaningful to the learner to carry out set tasks that have a purpose other than just the practice of language, i.e., tasks that reflect real-life activities. 3D multiuser virtual environments like Second Life provide a way to simulate immersive realistic scenarios that require meaningful communication within the framework of a formal curriculum and the constraints of the foreign language classroom. The lessons conducted on Chinese Island have been designed to be a blend of content taught in the main curriculum and textbook and real-life scenarios. Lesson scenarios include visiting a virtual restaurant, buying ingredients to make a traditional Chinese dish in a virtual farmers market, seeing a doctor, buying railway tickets, and renting accommodation among others.

5.2.5 Non-Player Characters

In the process of completing specific tasks set for each lesson/scenario, students communicate with virtual characters (non-player characters or NPCs) in Chinese in typed text chat. The NPCs, who play specific roles in each scenario, respond to student input in both Chinese text and audio. The tasks are designed so that the only way students are able to complete them is through communication with the NPCs in Chinese. All communication is in natural language rather than via set menus and conversation trees but is guided by specific information or actions that students need to obtain or carry out to complete each task.

It is important to point out that the Chinese Island NPCs are not simply chatbots. They perform a number of important roles that help make the interaction more authentic, e.g., provide linguistic interaction, provide feedback, and NoM/NfM (Negotiation of Meaning/Negotiation for Meaning which include recasts, clarification requests, repetition, confirmation, comprehension checks), perform tasks such as accepting payment and handing over goods, act as gatekeepers for the different stages of the tasks set. In addition to being able to conduct conversational interaction with learners, NPCs perform specific roles (e.g., a waitress, a market seller, a doctor

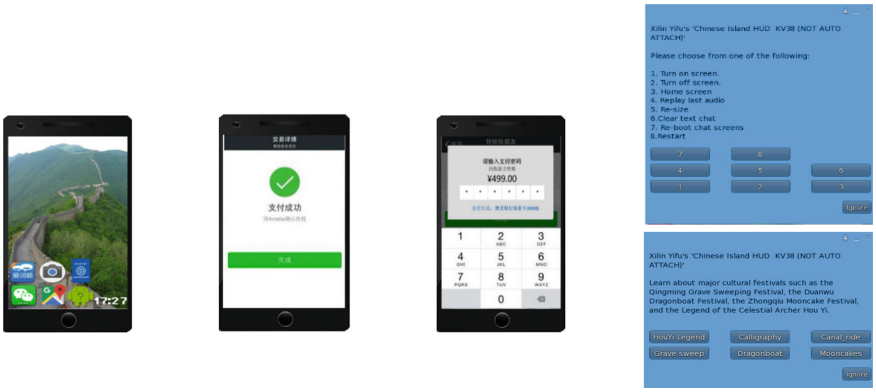


Fig. 5.1 Chinese Island HUD

in a hospital) within the specific scenario. They need to give the impression of being another "person" within the scenario. They also need to perform other functions other than simply "chatting" (e.g., initiating and accepting payment, and then handing over the item paid for in the medical scenario). The Chinese Island NPCs constitute a form of simple AI. Their chat engine is based on a set up similar to what was described earlier with a relatively simple pattern-matching AI.

One of the advantages of using NPCs is that they can interact with multiple students simultaneously without getting tired or confused. This frees up the teacher to interact with individual students as problems/questions arise, making the lessons student-focused with the teacher as facilitator. Students also interact with the NPCs independently without interfering with each other. This is achieved through an additional purpose designed in-world interface called the Chinese Island heads up display (HUD). The HUD is in the form of a virtual mobile phone and acts as a channel via which the NPC audio and text are transmitted to each student individually. The HUD also has other functions such as enabling the simulation of payment for items via an interface on the virtual phone called MeChat (an obvious play on WeChat which is now one of the main ways people in mainland China pay for things) (Fig. 5.1).

5.2.6 Lesson Structure, Instructions, and Assessment

As mentioned above, the lessons tasks are based on content covered in the main textbook used throughout the curriculum with the addition of extra content from real-life scenarios. The structure adopted for the lessons is a conventional task-based learning one with three main stages: pre-task, task, and post-task. The specific structure can be seen in Fig. 5.2.

To maintain immersion in the virtual environment as much as possible, a Lesson Instructions HUD (heads up display) has been designed that contains all of the

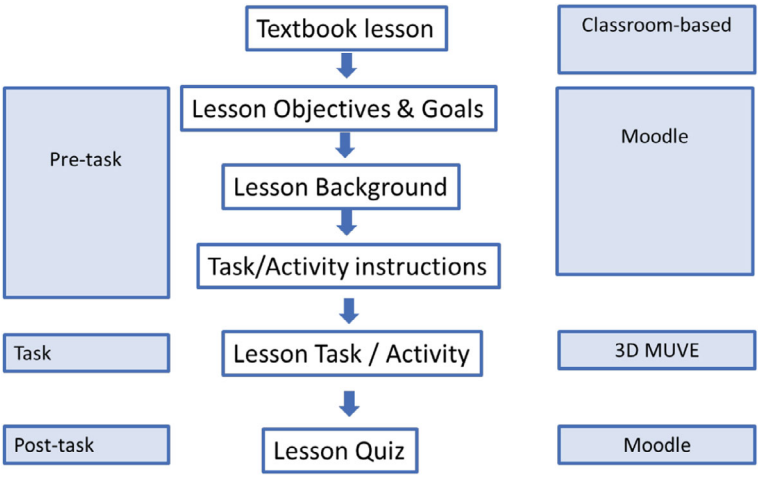


Fig. 5.2 Lesson structure

instructions related to the specific lesson/task the students are engaged in laid out in steps. The HUD is visible and can be interacted with from within the virtual environment and reduces the need for students to refer to materials outside of the virtual environment which can reduce the feeling of immersion. As each lesson/task involves assessment, the relevant steps of the Lesson Instructions HUD also contain links that open a quiz on Moodle specific to that part of the lesson/task (Fig. 5.3).

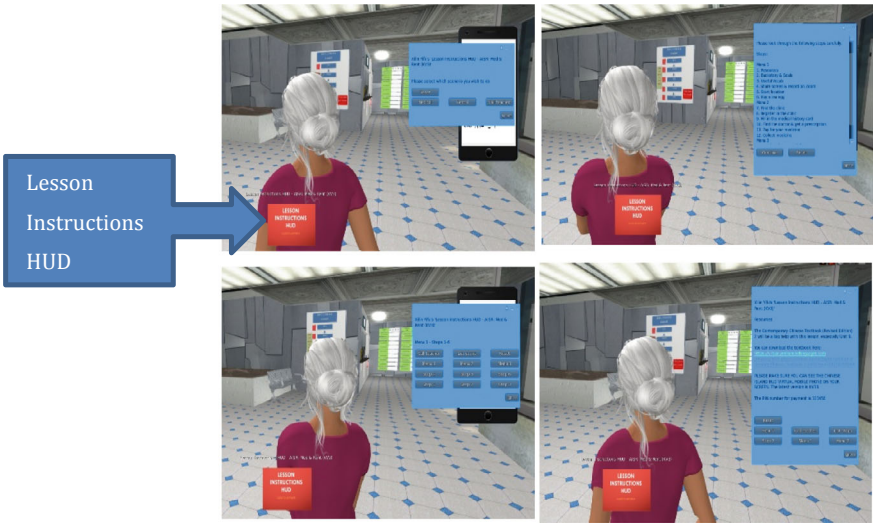


Fig. 5.3 Lesson Instructions HUD

5.3 The Pilot Study

Since 2008, over 2000 students have completed the lessons on Chinese Island. For the majority of this time, students have interacted with the NPCs via text chat in Chinese. Given the rapid advancement in recent years of technologies such as ASR (automatic speech recognition), in 2021, it was decided to conduct a series of studies involving the incorporation of ASR into the Chinese Island lessons to create an environment that can provide Chinese as a foreign language (CFL) learners with oral pragmatic practice.

The first stage involved the implementation of a pilot study to look at how ASR could be implemented in the Second Life/Chinese Island lesson environment and to gain insights from study participants on usability and usefulness of this kind of spoken interaction with the NPCs in terms of interactive oral practice, both in its own right and in comparison with typed interaction. Data to facilitate the harvesting of insights was gathered from an online questionnaire completed after each round of participation, videos recorded by the participants themselves, and logs of participant conversational interaction with the NPCs in the scenario.

The overarching research question of the pilot study is: How do the learners' reactions compare when using text-chat vs. voice chat to converse with NPCs? In particular, the following sub-questions are investigated:

- Benefits of each type of chat
- Ease of speaking with voice chat
- Ability of NPC to understand learners
- Ease of completing tasks
- Enjoyment of practicing with NPCs

5.3.1 Participants

Seven undergraduate students participated in the pilot study. Unfortunately, one only participated in the first part of the study and did not turn up for the second and so is not included in the following sections. All were students engaged in the ongoing study of Chinese language over a number of years within the Chinese Studies program of the host university. Their language levels ranged from CEFR (the Common European Framework of Reference for Languages) A2 to B2 (CEFR, 2022). In terms of the respective language levels at the host university, this equates to intermediate to advanced level Chinese. Two of the participants were male, four were female. Three of the participants were Australians with Chinese heritage, one an Australian with Vietnamese heritage, one an Australian with no Chinese heritage, and one of Vietnamese nationality.

5.3.2 The Study Design

The study involved participants completing one of the regular Chinese Island lessons conducted every year in a beginner level Chinese language unit in second semester twice: once using text chat to communicate with the NPCs, the second time using ASR and spoken interaction. At the end of each round, the participants filled out a 20-question survey that required them to respond to a range of statements using a scale of 1–5 to express how much they agreed or disagreed with the statement. After each question, they had the opportunity to provide open-ended feedback, and at the end of the survey, they were able to make any final comments about their experience. The questions can be loosely categorised into the following groups: benefits of conversing with the NPCs for Chinese speaking skills; comfort and ease of speaking with the NPCs; the NPCs' understanding and response; ease of completing tasks and understanding instructions; enjoyment of practicing speaking with the NPCs.

5.3.3 The Lesson Task

The specific lesson involved in the study required participants to visit a virtual medical clinic to see a doctor about either a bad cough or a stomach ache and to obtain medication to treat their symptoms. Completing the lesson task involved participants interacting with three different NPCs playing the roles of a nurse at reception, a doctor in the outpatient department, and a pharmacist in the clinic's internal pharmacy. When conducted as a normal part of the beginner level curriculum, in addition to providing students with the opportunity to practice using relevant vocabulary, grammar, and syntactic knowledge learned in the classroom in a conversational manner, a key goal is for students to also learn about the process of seeing a doctor in China. Generally, the process involves paying a registration fee at reception, receiving, and filling out a medical history card, showing the card to a doctor and then discussing symptoms with them, receiving a prescription for medication from the doctor, taking the prescription back to reception and paying for the medication, and then finally showing the payment receipt to the pharmacist in the clinic and being given the medication. Each of these steps requires talking to the relevant NPC to find out what needs to be done and in what order. The steps cannot be completed out of sequence or without communicating with the relevant NPC.

5.3.4 ASR Technology Used

For the purposes of the study, students were asked to install and activate the ASR system native to the operating system of the personal computer they were using. For

those using Windows PCs, Windows Speech Recognition, which is available natively from Windows Vista onwards, was used (Microsoft, 2022). Participants using Apple Mac computers used Dictate, available from MacOS Mojave 10.4 onwards (Apple, 2022). One of the major benefits of using these two ASR programs is that they are part of the operating system software, do not require any additional software to be installed (other than the relevant language package) and are easily activated. Moreover, both programs are able to be used with the Second Life interface. After clicking in the chat input field on the interface, users turn on their microphone and speak causing the ASR program to convert their spoken word into text which is then automatically input into the chat input field. A range of other commercially available ASR software looked at prior to the study was not able to output text directly into the Second Life interface chat input field. A second big advantage of these two native ASR programs is that they are free. Using commercial ASR programs is too expensive as a copy has to be bought for each student. Given that during normal lessons, there can be up to 100 participants, this would be prohibitively expensive.

While both ASR systems have a fairly high level of accuracy in terms of understanding what is being spoken, even with very standard pronunciation recognition accuracy is not 100%. Moreover, if pronunciation deviates from the standard form, the level of recognition accuracy decreases even further. As was found in this study, pronunciation can deviate for a number of reasons. Students with a lower level of linguistic skills and experience often are still in the process of mastering pronunciation. Even non-native speaking students at a much higher level can still have issues with pronunciation. For native speakers, the situation, particularly for Chinese, is more complex. Many native speakers fluent in Mandarin have non-standard regional accents influenced by their native dialects, which poses significant recognition problems for both the Windows and Apple ASR programs (this was one problem raised in participant feedback). There are ways to enable these ASR programs to better recognise regional accents which were not employed in this study (authors unaware of these options) but are recommended for future studies.

5.3.5 Pilot Study Results: Problems Encountered by Study Participants

Based on feedback from the participants, observation of the videos recorded by the participants of themselves completing the set task in the virtual environment, and detailed reading of the conversational logs, apart from initial set up problems the main issues encountered by participants were:

- *ASR related*
- The ASR program often recognised words they spoke incorrectly and thus output incorrect characters.
- This led to participants having to repeat a word or phrase multiple times until the desired Chinese characters were output by the ASR program.

- Operational difficulties. Participants sometimes forgot to click in the Second Life chat input field first before activating the speech recognition function and commencing speaking.
- Some students forgot to activate Chinese language input before commencing speaking, causing the speech recognition system to be unable to understand their spoken output.
- *Communicating with the NPCs*
- NPC responses were quite limited, and in some cases required fairly specific input to achieve an appropriate response from the NPC. Where an NPC did not understand the input from a participant, it did not offer any guidance as to why it did not understand leaving the participants to work out what the problem was themselves.
- Because the content and vocabulary the NPCs are programmed to understand was set for beginner level students, when participants used more advanced ways of expressing things this sometimes led to incomprehension on the part of the NPCs.

5.3.6 Student Preferences for Either Spoken or Typed Conversations

In terms of rating their preference for voice-based or text-based interaction with the NPCs, three participants rated speaking as being better than typing (one marginally, two more strongly), while two preferred typing (one marginally, one slightly stronger).

Reasons for preferring speaking to typing included:

> "I liked speaking better than typing as it engaged me more…. It was very comfortable and helped practise tones. It could be a bit slow fiddling around with the dictation, then switching to the keyboard but it felt just as rewarding as typing." (Female, Vietnamese Australian, CEFR A2, ASR: macOS)

> "It was more useful to my learning of the words and pronunciation when I need to say it out loud rather than typing the pinyin or copy and pasting. The voice function was easy for me to set up and activate whenever I needed it." (Female, Chinese heritage Australian, CEFR B1, ASR: Windows)

Some qualified comments about preferring speaking to typing were also made such as:

> "I think that the combination of speaking with the voice recognition technology really enhances the experience with the NPCs as it feels like you're actually talking to someone however if you're like me and struggle to pronounce some words, it's quite the tedious and long session." (Male, Chinese heritage Australian, CEFR A1+)

> "Typing is a lot easier than speaking, due to pronunciation issue I had last time however I do reckon speaking is a lot more fun than typing. Although a mixture of the two wouldn't be all that bad either." (Male, Australian, CEFR B1)

Reasons for preferring typing to speaking included:

> "The dictation on my MAC hardly picks up what I want to say and I have to type them in sometimes. I might pronounce words incorrectly so the AI cannot recognise it." (Female, Chinese heritage Australian, CEFR A1+, ASR: macOS).
>
> "I feel that this would be a really fun way to practise chinese however I do believe it's quite time consuming, especially coming up with your own sentences, and pronouncing them correctly so that the computer knows what you intended to say." (Male, Chinese heritage Australian, CEFR A1+).

5.4 The Larger Study

The results and insights from the pilot study will inform the design and implementation of a larger study that will investigate the following research questions:

- Research question 1: What are the design principles and usability of chatbots that combine AI speech recognition in CFL oral interactive practice?
- Research question 2: What are the design principles of the CFL oral interactive practice environment made up of AI speech recognition and VR, and what is the usability in judging and providing feedback on the oral output of CFL learners?
- Research question 3: What is the effect of the CFL oral interactive practice environment made up of AI speech recognition and VR on the attitude of CFL learners to oral communication?
- Research question 4: What is the impact of the CFL oral language interactive practice environment combined with AI speech recognition and VR on the oral pragmatic ability of CFL learners?

5.4.1 Participants and Execution

While in the pilot study participants came from across a range of different classes and language levels, in the planned larger study all participants will come from one beginner level class. Moreover, the study will be conducted as part of two a regular scheduled lessons on Chinese Island late in the first semester after students have been attending classes for 5 weeks (the first time) and 8 weeks (the second time). The content and difficulty of the lessons will reflect the content learned in the classroom in the preceding weeks.

5.4.2 Set up Problems

In addition to comments in participant feedback about the speech recognition program not picking up their voices clearly and thus not recognising what was being said, anecdotal comments made by participants in the lead up to the pilot study indicated that some participants has difficulty setting up and activating either Windows

Speech Recognition or macOS Dictation on their computers. While detailed written information was provided to participants several weeks before the pilot study, providing additional video explanations demonstrating step-by-step what is required and then setting practice examples for participants to complete in the lead up to a larger study should improve the situation. Initially these short practice examples could be done in a Word document with set texts (to remove the cognitive burden of having to think what to say) so that the focus can be placed on setting up appropriate volume levels and ensuring that speech recognition is working as expected. Where necessary, participants who have particular difficulty setting up speech recognition can make appointments to meet with a technical advisor to solve any outstanding problems.

5.4.3 ASR Recognition Accuracy

A number of measures could be considered to increase recognition accuracy. Since the time of the pilot study (late 2021), more advanced speech recognition systems with a greater ability to handle non-standard pronunciation have become more readily available. These include online services like Microsoft Azure AI Speech (https://azure.microsoft.com/en-us/products/ai-services/ai-speech), Amazon Transcribe (http://tinyurl.com/Amazon-Transcribe), and Open.ai Whisper (https://openai.com/research/whisper), as well as downloadable programs like Dragon Speech Recognition (https://www.nuance.com/en-au/dragon.html). However, as each of these services requires either or a cost for their use or a relatively high level of technical expertise to set up, and they are not guaranteed to work with the Second Life interface, they will not be considered.

As with the pilot study, the larger study will continue to use the native Windows Speech Recognition and macOS Dictation programs installed on Windows and Mac computers due to these applications being free and relatively easy to get set up and running. They have also been shown to work well with the Second Life interface. The main solution to improve recognition accuracy is for participants to use the inbuilt functionality of both these programs to train them in better recognition of the user's voice. The completion of the above-mentioned short practice examples in the weeks leading up to the study will provide an opportunity for participants to carry out this accuracy training.

5.4.4 Feedback on Pronunciation

One of the limitations of the Windows and macOS speech recognition systems is the lack of feedback in relation to pronunciation accuracy and correction. As indicated by participants in the pilot study, when their pronunciation is not standard the characters produced by the speech recognition system were often not what was intended. In a

sense, this is one form of feedback, albeit not in terms of providing a guide to correct pronunciation. In an attempt to provide some kind of feedback on correct pronunciation, participants in the large study will be encouraged to use Google Translate. After manually typing the character, word, or phrase they wish to speak into Google Translate, they can then listen to the correct pronunciation, practice it for a few moments, then try speaking to the speech recognition system again. In order to reduce the impact on immersion in the virtual environment, access to Google Translate will be provided via the Second Life interface.

5.4.5 NPC Responses and Additional Feedback

To overcome the issue identified by participants in the pilot study where some responses received from the NPCs were not instructive or helpful due to the fact that the NPCs were unable to understand participant input, a functionality has been added to the NPCs where anything they do not understand will be sent to ChatGPT for grammatical, semantic, and syntactical analysis in relation to the specific context of the scenario. This analysis will be automatically sent back to the participants within the virtual environment and will hopefully provide some guidance on what the problem was and how to rectify it. Recent ad hoc experiments with this approach have even elicited pronunciation correction (via Romanised pinyin spelling) where ChatGPT has divined what the speaker intended to say from the scenario context and the context of their overall utterance.

5.5 Conclusion

Both the literature and feedback from participants in the pilot study indicate that in addition to the benefits of interacting with virtual characters (NPCs) in appropriately designed virtual environments, being able to interact with them in voice has additional potential benefits over text-based interactions. These include increased motivation (fun) for some, ease of use compared to typing (for some), the opportunity to practice pronunciation, and the opportunity to communicate in a manner closer to that of a real conversation. With the rapid development of powerful AI and adjacent technology (speech recognition, artificial voices, etc.) in recent years, voice-based communication with embodied AI will become an even greater part of our everyday lives in the future. As the cost of these technologies comes down, their use as an additional tool in the language educator's and language learner's tool box will become more viable and more and more common. The full study planned will provide more detailed and systematic empirical evidence of the usability and impact of voice-based interaction with chatbots and feedback provided by ASR.

References

Allen, H. W., & Herron, C. (2003). A mixed-methodology investigation of the linguistic and affective outcomes of summer study abroad. *Foreign Language Annals, 36*(3), 370–385.

Apple. (2022). Dictate messages and documents on Mac. In *macOS User Guide.* https://support.apple.com/en-au/guide/mac-help/mh40584/mac

Australian Bureau of Statistics. (2022). Cultural diversity: Census. Information on country of birth, year of arrival, ancestry, language, and religion. https://www.abs.gov.au/statistics/people/people-and-communities/cultural-diversity-census/2021

Beauvois, M. (1997). Write to speak: The effects of electronic communication on the oral achievement of fourth-semester French students. In J. A. Muyskens (Ed.), *New ways of learning and teaching: Focus on technology and foreign language education* (pp. 93–116). Heinle & Heinle.

Bojic, L. (2022). Metaverse through the prism of power and addiction: What will happen when the virtual world becomes more attractive than reality? *The European Journal of Futures Research, 10*(22), 1. https://doi.org/10.1186/s40309-022-00208-4

Chandra, A.L. (2018, July 20). Mankind's First Mathematical Model Of A Biological Neuron. *Medium.* https://towardsdatascience.com/mcculloch-pitts-model-5fdf65ac5dd1

Ciampelli, S., Voppel, A. E., de Boer, J. N., Koops, S., & Sommer, I. E. C. (2023). Combining automatic speech recognition with semantic natural language processing in schizophrenia. *Psychiatry Research, 325*, Article 115252. https://doi.org/10.1016/j.psychres.2023.115252

De Clercq, G. (2023, January 28). Top French university bans use of ChatGPT to prevent plagiarism. *Reuters.* https://www.reuters.com/technology/top-french-university-bans-use-chatgpt-prevent-plagiarism-2023-01-27/

Collentine, J. (2009). Study abroad research: Findings, implications, and future directions. In *The Handbook of Language Teaching* (pp. 218–233). Wiley.

Collier, K. (2023, May 31). AI risks leading humanity to 'extinction,' experts warn. *NBC News.* https://www.nbcnews.com/tech/tech-news/ai-risks-leading-humanity-extinction-experts-warn-rcna86791

Collins, L., & Muñoz, C. (2016). The foreign language classroom: Current perspectives and future considerations. *The Modern Language Journal, 100*(S1), 133–147.

Common European Framework of Reference for Languages (CEFR). (2022). https://www.coe.int/en/web/common-european-framework-reference-languages/table-2-cefr-3.3-common-reference-levels-self-assessment-grid

Cordella, M., & Huang, H. (2016). *Rethinking second language learning: Using intergenerational community resources.* Multilingual Matters.

Davis, A. (2023, January 30). ChatGPT banned in WA public schools in time for start of school year. *ABC Australia.* https://www.abc.net.au/news/2023-01-30/chatgpt-to-be-banned-from-wa-public-schools-amid-cheating-fears/101905616

Davis, A. (2023, January 26). ChatGPT sparks cheating, ethical concerns as students try realistic essay writing technology. *ABC Australia.* https://www.abc.net.au/news/2023-01-26/chatgpt-sparks-cheating-ethical-concerns-in-schools-universities/101888440

De Back, T. T., Tinga, A. M., & Louwerse, M. M. (2023). Learning in immersed collaborative virtual environments: Design and implementation. *Interactive Learning Environments, 31*(8), 5364–5382. https://doi.org/10.1080/10494820.2021.2006238

Deighan, M. T., Ayobi, A., & O'Kane, A. A. (2023). Social virtual reality as a mental health tool: How people use VRChat to support social connectedness and wellbeing. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23)*, April 23–28, 2023, Hamburg, Germany. ACM, New York, NY, USA, 18 Pages. https://doi.org/10.1145/3544548.3581103

Dictate.it. (2021, March 16). Speech Recognition from Audrey to Alexa—A Brief History. Retrieved January 18, 2024, from https://dictateit.com/speech-recognition-from-audrey-to-alexa-a-brief-history/

Dsouza, R., Sahu, S., Patil, R., Kalbande, D. R. (2019). Chat with bots intelligently: A critical review analysis. In *2019 6th IEEE International Conference on Advances in Computing, Communication and Control (ICAC3 2019).* Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.1109/ICAC347590.2019.9036844

Dundas, K. (2021). What is the metaverse. *Techradar.* https://www.techradar.com/news/the-metaverse-is-coming-but-what-does-that-even-mean

Dwivedi, Y. K., Hughes, L., Baabdullah, A. M., Ribeiro-Navarrete, S., Giannakis, M., Al-Debei, M. M., Dennehy, D., Metri, B., Buhalis, D., Cheung, C. M. K., Conboy, K., Doyle, R., Dubey, R., Dutot, V., Felix, R., Goyal, D. P., Gustafsson, A., Hinsch, C., Jebabli, I., … Wamba, S. F. (2022). Metaverse beyond the hype: Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. *International Journal of Information Management, 66*, Article 102542. https://doi.org/10.1016/j.ijinfomgt.2022.102542

Erickson, T. (1999). Persistent conversation: An introduction. *Journal of Computer-Mediated Communication, 4*(4), 1.

Falvey, M. (1988). Classroom Interaction Research and the Foreign Language Classroom. *Working Papers in Linguistics and Language Teaching, 1988*(11), 18.

Interaction Design Foundation - IxDF. (2019, February 26). A Brief History of Human-Computer Interaction. *Interaction Design Foundation - IxDF.* https://www.interaction-design.org/literature/article/a-brief-history-of-human-computer-interaction

Garrett, U. (2023, August 11). What is Roblox? Here's everything you need to know. *CNN.* https://edition.cnn.com/cnn-underscored/electronics/what-is-roblox

Gass, S. M., & Selinker, L. (2008). *Second language acquisition: An introductory course* (3rd ed.). Routledge.

Golonka, E. M., Bowles, A. R., Frank, V. M., Richardson, D. L., & Freynik, S. (2014). Technologies for foreign language learning: A review of technology types and their effectiveness. *Computer Assisted Language Learning, 27*(1), 70–105. https://doi.org/10.1080/09588221.2012.700315

Gregory, A. & Hern, A. (2023, May 10). AI poses existential threat and risk to health of millions, experts warn. *The Guardian.* https://www.theguardian.com/technology/2023/may/10/ai-poses-existential-threat-and-risk-to-health-of-millions-experts-warn

Hawley, M. (2023, November 3). Generative AI Timeline: 9 Decades of Notable Milestones. *CMS Wire.* Retrieved on January 16, 2024, from https://www.cmswire.com/digital-experience/generative-ai-timeline-9-decades-of-notable-milestones/

Hunt, S. (2023, May 21). "Why Mark Zuckerberg's metaverse gamble failed". *The Spectator.* https://thespectator.com/topic/mark-zuckerberg-metaverse-gamble-failed-meta/#:~:text=So%20Zuck%20felt%20forced%20to,still%20recuperating%20from%20Zoom%20fatigue

Ibrahim, D. (2023, February 14). What is the Fortnite Metaverse: Reality or Fiction. *The Metaverse Insider.* https://metaverseinsider.tech/2023/02/14/fortinte-metaverse/

James, A. (2023, March 30). ChatGPT has passed the Turing test and if you're freaked out, you're not alone. *TechRadar.* https://www.techradar.com/opinion/chatgpt-has-passed-the-turing-test-and-if-youre-freaked-out-youre-not-alone

Jones, K. S. (1994). Natural Language Processing: A Historical Review. In A. Zampolli, N. Calzolari, & M. Palmer (Eds.), *Current issues in computational linguistics: In honour of Don Walker* (Linguistica Computazionale, vol. 9–10, pp. Pisa, Dordrecht). https://aclanthology.org/www.mt-archive.info/Zampolli-1994-Sparck-Jones.pdf

Khan, S. (2023, August 23). Machine Learning Breakthroughs: Transforming the Landscape of Artificial Intelligence. *Medium.* https://medium.com/@fdkhan2587/machine-learning-breakthroughs-transforming-the-landscape-of-artificial-intelligence-7e70131dba9a

Kodish-Wachs J, Agassi E, Kenny P 3rd, Overhage JM. (2018). A systematic comparison of contemporary automatic speech recognition engines for conversational clinical speech. In *AMIA Annu Symp Proc* (pp. 683–689). PMID: 30815110; PMCID: PMC6371385.

Li, J. (2014, September). Deep Neural Network for Automatic Speech Recognition: from the Industry's View [PowerPoint slides]. *Microsoft.* https://www.microsoft.com/en-us/research/wp-content/uploads/2014/09/NTU_2014.pdf

Llanes, A. (2011). The many faces of study abroad: an update on the research on L2 gains emerged during a study abroad experience. *International Journal of Multilingualism, 8*(3), 1.

Mac Ghlionn, J. (2023, February 25). The metaverse has failed? It never even took off. *South China Morning Post.* https://www.scmp.com/comment/opinion/article/3211121/metaverse-has-failed-it-never-even-took

Maddock, B. (2023, October 18). *Insider Gaming.* https://insider-gaming.com/5-best-fortnite-live-events-of-all-time/

Manaris, B. (1998). Natural Language Processing: A Human-Computer Interaction Perspective. In M. V. Zelkowitz (Ed.), *Advances in Computers* (Vol. 47, pp. 1–66). Elsevier. https://doi.org/10.1016/S0065-2458(08)60665-8

Manavis, S. (2023, May 31). The death of the Metaverse. *The New Statesman.* https://www.newstatesman.com/culture/social-media/2023/05/death-metaverse-mark-zuckerberg-facebook

Marr, B. (2023, May 19). A Short History Of ChatGPT: How We Got To Where We Are Today. *Forbes.* https://www.forbes.com/sites/bernardmarr/2023/05/19/a-short-history-of-chatgpt-how-we-got-to-where-we-are-today/?sh=f441a84674f1

Microsoft. (2022). Use voice recognition on Windows. https://support.microsoft.com/en-us/windows/use-voice-recognition-in-windows-83ff75bd-63eb-0b6c-18d4-6fae94050571#WindowsVersion=Windows_11

Milmo, D. (2021). Enter the metaverse: the digital future Mark Zuckerberg is steering us toward. *The Guardian.* https://www.theguardian.com/technology/2021/oct/28/facebook-mark-zuckerberg-meta-metaverse?CMP=share_btn_link

CDC Museum. (n.d.). *COVID-19 Timeline.* https://www.cdc.gov/museum/timeline/covid19.html

Pandey, S., & Sharma, S. (2023). A comparative study of retrieval-based and generative-based chatbots using Deep Learning and Machine Learning. *Healthcare Analytics, 3*, 100198. https://doi.org/10.1016/j.health.2023.100198

Parkin, S. (2023, June 10). Who needs the Metaverse? Meet the people still living on Second Life. *The Guardian.* https://www.theguardian.com/technology/2023/jun/10/who-needs-the-metaverse-meet-the-people-still-living-on-second-life

Peters, J. (2023, December 15). Roblox: all the news about the popular social and gaming platform. *The Verge.* https://www.theverge.com/23864790/roblox-news-announcements-updates/archives/2

Powers, D.M.W. (1998). The total Turing test and the Loebner prize. In *Proceedings of the Joint Conferences on New Methods in Language Processing and Computational Natural Language Learning (NeMLaP3/CoNLL '98).* Association for Computational Linguistics, USA, 279–280.

Reid, K. (2023, July 5). Are you listening to me? The inner workings of speech recognition. *Australian National University School of Cybernetics.* https://cybernetics.anu.edu.au/news/2023/07/05/are-you-listening-to-me/

Reimers, F. M. (2022). Learning from a pandemic. The impact of COVID-19 on education around the world. *Primary and secondary education during Covid-19: Disruptions to educational opportunity during a pandemic* (pp. 1–37).

Rudra, S. (2023, March 30). ChatGPT in Education: The Pros, Cons and Unknowns of Generative AI. *EdTech.* https://edtechmagazine.com/k12/article/2023/03/chatgpt-education-generative-ai-perfcon

Sands, R. A. (2006). Comparison of Chat Room Productivity: In-class Versus Out-of-class. *CALICO Journal, 24*(1), 59–76.

Simon Says (2023). What is ASR? How Does it Work? Our In-Depth 2023 Guide. January 18, 2024, from https://www.simonsaysai.com/blog/automatic-speech-recognition

Sheiber, S. (2011, November). Inverting the Turing Test. *American Scientist.* https://www.americanscientist.org/article/inverting-the-turing-test

Somoye, F.L. (2023, September 4). Can ChatGPT pass the Turing Test? *PC Guide.* https://www.pcguide.com/apps/chat-gpt-pass-turing-test/#:~:text=However%2C%20according%20to%20Metaverse%20Post,cases%2C%20managed%20to%20deceive%20evaluators

Standish, J. (2023, May 25). What Generative AI Means For Retail. *Forbes.* https://www.forbes.com/sites/jillstandish/2023/05/25/what-generative-ai-means-for-retail/?sh=4a4d9dd768e2

Sun, W. (2023, August 16). The impact of automatic speech recognition technology on second language pronunciation and speaking skills of EFL learners: a mixed methods investigation. *Front Psychol, 14*, 1210187. https://doi.org/10.3389/fpsyg.2023.1210187

Taguchi, N. (2015). "Contextually" speaking: A survey of pragmatic learning abroad, in class, and online. *System, 48*, 3–20.

Taguchi, N. (2018). Contexts and pragmatics learning: Problems and opportunities of the study abroad research. *Language Teaching, 51*(1), 124–137.

Tarnoff, B. (2023, July 23). Weizenbaum's nightmares: how the inventor of the first chatbot turned against AI. *The Guardian.* https://www.theguardian.com/technology/2023/jul/25/joseph-weizenbaum-inventor-eliza-chatbot-turned-against-artificial-intelligence-ai#:~:text=In%201966%2C%20an%20MIT,typewriter%20connected%20to%20a%20mainframe

Taulli, T. (2023, May 30). How Generative AI Will Remake Cybersecurity. *eSecurity Planet.* https://www.esecurityplanet.com/trends/generative-ai-cybersecurity/

Thorne, S. (2000). Beyond Bounded Activity Systems: Heterogeneous Cultures in Instructions Uses of Persistent Conversation. Paper presented at the 33rd Hawaii International Conference on System Sciences, Hawaii.

Thorne, S. (2008). Mediating Technologies and Second Language Learning. In J. Coiro, M. Knobel, C. Lankshear, & D. J. Leu (Eds.), *Handbook of Research on New Literacies* (pp. 415–448). Lawrence Erlbaum.

Toloka.ai (2023, August 22). History of generative AI. January 16, 2024, from https://toloka.ai/blog/history-of-generative-ai/

Tyagi, N., & Bhushan, B. (2023). Demystifying the role of natural language processing (NLP) in smart city applications: Background, motivation, recent advances, and future research directions. *Wireless Personal Communications, 130*, 857–908. https://doi.org/10.1007/s11277-023-10312-8

Valance, C. (2023, March 28). AI could replace equivalent of 300 million jobs—report. *BBC News.* https://www.bbc.com/news/technology-65102150

Vanian, J., & Levy, A. (2023, February 1). Meta lost $13.7 billion on Reality Labs in 2022 as Zuckerberg's metaverse bet gets pricier. *CNBC.* https://www.cnbc.com/2023/02/01/meta-lost-13point7-billion-on-reality-labs-in-2022-after-metaverse-pivot.html

Vanian, J., Goswani, R. (2023, April 19). Meta has started its latest round of layoffs, focusing on technical employees. *CNBC.* https://www.cnbc.com/2023/04/19/meta-started-latest-round-of-layoffs-focusing-on-technical-employees.html

Wagner, J.A. (2023b). Watch: VRChat Deaf Community Creates New Dialect Of Sign Language To Accommodate Limited VR Hand Tracking. *New World Notes.* https://nwn.blogs.com/nwn/2023/10/vrchat-new-dialect-of-asl.html

Wagner, J.A. (2023a). Second Life's Human Avatars Have Become Amazingly Ultra-Realistic Over The Years -- Even If That Hasn't Grown The User Base At All. *New World Notes.* https://nwn.blogs.com/nwn/2023/10/hyper-real-avatars-zuckerberg-meta-second-life.html

Wei, J. (2019, July 3). AlexNet: The Architecture that Challenged CNNs. *Medium.* https://towardsdatascience.com/alexnet-the-architecture-that-challenged-cnns-e406d5297951

Zamora, G. (2023, April 29). *PCMag Australia.* A Beginner's Guide to Fortnite: 12 Tips for your First Match. https://au.pcmag.com/features/57811/a-beginners-guide-to-fortnite-12-tips-for-your-first-match

Dr. Scott Grant lived studied and worked in China for 8 years. He has taught Chinese language and culture at tertiary level for more than 20 years. He coordinates and teaches two gateway units and two capstone units in the Monash Chinese Studies program. He is the creator of the Monash Chinese Island virtual language learning simulation in Second Life that has been part of

the formal introductory level curriculum at Monash since 2008. Scott has also conducted collaborative research and jointly published a number of papers on educational uses of 3D multi-user virtual environments covering topics including self-efficacy (自我效能感), cognitive skills (认知能力), second language acquisition (第二种语言习得) and foreign language anxiety (外语焦虑). His PhD thesis was based on his work on task-based language learning in the virtual environment of Chinese Island. Scott's current research focuses on educational uses of high immersion VR and AI.

Chapter 6
Applying AI Technologies in Second Language Learning

Hao-Jan Chen, Fu-Chan Hsieh, and Yu-Ci Yen

Abstract This paper explores the evolution and impact of Artificial Intelligence (AI) technologies in second language learning, particularly focusing on speaking and writing skills. It traces the development from early Computer-Assisted Language Learning (CALL) to current AI-powered tools, highlighting the transformative potential of technologies like Automatic Speech Recognition (ASR), Text-to-Speech (TTS), Intelligent Personal Assistants (IPAs), and AI-driven chatbots. The paper discusses various applications, including Cool English's speech assessment tool, CoolE Bot for conversational practice, and AI-based grammar checkers. While acknowledging the benefits of these technologies in providing personalized learning experiences and immediate feedback, the paper also addresses challenges such as potential biases, ethical concerns, and the need for critical evaluation of AI-generated content. The authors emphasize the importance of balancing AI integration with traditional teaching methods and fostering responsible use of these tools in language education.

Keywords Artificial intelligence in language learning · Automated speech recognition (ASR) · Text-to-speech (TTS) · AI-powered Chatbots · Automated writing evaluation (AWE)

6.1 Introduction

Over the past decades, there has been substantial transformation in second language acquisition, which is largely driven by technological advancement. Computer-Assisted Language Learning (CALL) marked a crucial juncture, initiating a new era where technology intersected with pedagogical practices, and introducing innovative pathways to enhance language learning. Previous CALL initiatives have introduced an array of computer-based tools and resources aimed at complementing

H.-J. Chen (✉) · F.-C. Hsieh · Y.-C. Yen
English Department, National Taiwan Normal University, Taipei City, Taiwan
e-mail: hjchenntnnu@gmail.com
URL: http://www.coolenglish.edu.tw

Y. Lan et al. (eds.), *AI-Mediated Language Education in the Metaverse Era*, Chinese Language Learning Sciences, https://doi.org/10.1007/978-981-95-0245-5_6

traditional classroom teaching. These pioneering programs, employing interactive exercises, multimedia materials, and self-paced learning activities, have revolutionized conventional teaching methodologies by integrating technology into language learning processes. The emergence of OpenAI ChatGPT in 2022 represents a monumental leap in using technologies in second/foreign language education. In contrast to earlier CALL methods, AI-assisted language learning harnesses machine learning, natural language processing, and adaptive algorithms, fundamentally reshaping the language acquisition environment.

While language teaching has always been seen as a preparation for out-of-class uses of language, much of the focus in language teaching in the past has typically been on classroom-based language learning (Richards, 2015). As Internet technology progresses, it is convenient for students to effectively utilize diverse computer and internet resources to complement their English language learning during their leisure time. Gaining mastery in English goes beyond the confinement of classroom instruction; it is imperative to make efficient use of time beyond the structured learning environment. Technology-based activities offer rich opportunities for second language learning, and foster linguistic, communicative, and pragmatic competence (Zhao & Lai, 2023). Learners are expected to enhance accuracy and fluency, engage in extended language exposure, tap into multimodal learning sources, and develop autonomous learning skills. In addition, out-of-class activities might bring unique benefits to teachers by providing authentic learning opportunities that surpass traditional classroom settings. These activities facilitate the fusion of classroom and external learning opportunities and enable teachers to connect classroom learning with other real-world tasks.

In light of these benefits, it is suggested that teachers actively encourage and guide their students in utilizing technology outside the classroom to enhance students' language learning. Teachers' support comes in different forms, including sharing information about beneficial technologies and resources and offering guidance on their use (Lai & Gu, 2011). Regardless of the methods, the essential goal is for teachers to consciously prioritize this encouragement and make it an integral part to enable students to harness the advantages of technology in their language learning process. Nowadays, AI technologies hold immense potential to transform traditional instructional methods, providing personalized learning experiences tailored to individual needs and preferences (Hwang et al., 2020). The adaptability and personalization of AI-based learning breaks through the conventional CALL era, signifying a profound departure from more standardized instructional approaches toward more individualized, learner-centered methodologies.

The incorporation of AI technologies into language education goes beyond customized instruction, including adaptive feedback systems, instant evaluations, and immersive learning encounters. The integration of AI-driven chatbots, natural language processing (NLP) algorithms, and virtual tutors might transform the learning mode, fostering autonomy and self-directed learning among students outside of the classroom.

With the advancement in AI, several key players of automatic speech recognition (ASR) have significantly improved their ASR engines. Although ASR is not a

new technology, the accuracy and performance of the recent ASR systems are truly impressive. These ASR tools provide instantaneous corrective feedback to enable L2 learners to check their pronunciation and speaking skills independently, without requiring additional support from teachers (García et al., 2020). They patiently interact with learners, quickly and accurately identifying errors in their oral production. If learners can have access to these new tools, it is clear they can further notice their problems and improve based on the machine feedback.

A few years ago, Intelligent Personal Assistants (IPAs) experienced significant popularity, becoming widely utilized across various contexts. Both AI and NLP (Natural Language Processing) techniques play crucial roles in enabling (IPAs) to comprehend and generate human-like language, making interactions more intuitive and user-friendly. IPA used Automatic Speech Recognition (ASR) and various other NLP technologies to allow users to interact more freely with their voices. These tools are particularly attractive for second and foreign language learning, as they enable L2 learners to engage with IPAs in their target languages. By integrating speech recognition and intelligent language understanding, IPAs provide a more user-friendly interactive environment. This technology has practical applications in various fields, including personal assistants like Siri, Google Assistant, and Alexa on smart speakers and smartphones. IPAs also help learners easily access authentic language input anytime, anywhere, enhancing second and foreign language learning.

Even though assistants like Apple Siri, Google Assistant, and Amazon Alexa provide high-quality learning materials, they are more limited in their ability to interact freely with language learners. These assistants are not able to engage with L2 learners on a wide range of topics due to the limited data in their knowledge databases. Their restricted knowledge prevents them from responding effectively to unfamiliar subjects. For many years, researchers have aimed to develop high-quality chatbots for second and foreign language learners, believing that chatbots could serve as promising tools to enhance various language skills. In a review study, Huang et al. (2022) revealed that chatbots are not only effective in facilitating students' oral skills but also function as interlocutors to help students practice grammar (e.g., Kim et al., 2019; Wang et al., 2017) and assist with essay writing (e.g., Lin & Chang, 2020). Although these studies presented some positive findings, the general outcomes were still often challenged by students, teachers, and researchers. The core issue was the lack of an intelligent database capable of answering a wide range of student questions. However, with the emergence of Large Language Models (LLMs), such as Google Bard, OpenAI ChatGPT, and Claude 2, the intelligence of machine learning has significantly improved. AI tools based on LLMs have transformed the development of chatbots for language learning. Educators can now assume roles as facilitators and mentors, guiding students through a variety of AI-enhanced chatbots, tailoring instruction to individual learning styles, and leveraging technology to foster linguistic proficiency.

Another major breakthrough is the development of AI-based grammar checker. For many decades, various AWE systems are designed to analyze various aspects of writing, including vocabulary, grammar, register, and organization. Ngo et al. (2022) analyzed several key studies in regard to the effectiveness of Automated

writing evaluation (AWE) on EFL/ESL students' writing performance, and the results showed that AWE is more effective in improving vocabulary usage but less effective in improving grammar in students' writing. For many L2 learners, Grammarly's AWE systems have the potential to serve as an effective writing assistant (Dizon & Gayed, 2021; Fitria, 2021). However, with the emergence of ChatGPT, AI-based grammar checkers have begun to challenge the effectiveness of Grammarly, which relies on more traditional NLP technologies, potentially prompting Grammarly to integrate ChatGPT into its own grammar-checking system.

This chapter focuses on two major areas: speaking tools and writing tools. In the first section, we will explore various speaking tools, including Automatic Speech Recognition (ASR) systems, Text-to-Speech (TTS) systems, and smart speakers. We will also introduce Cool-EBot, a generative AI-powered chatbot developed by National Taiwan Normal University, which leverages Microsoft's Azure OpenAI Service to assist students in practicing English conversation and speaking skills. A notable feature of Cool-EBot is its integrated mini grammar checker, which helps students verify the accuracy of their spoken output. In the second section, we will turn to writing tools, highlighting the capabilities of ChatGPT-based grammar checker, designed to improve the quality of automated feedback on second language writing.

6.2 The Need for Quality Feedback on Speaking Activities

In the Asian ESL/EFL context, the primary linguistic learning objective is typically not focused on developing proficiency in spoken English within mainstream education. This is attributed to the considerable time investment required to teach English-speaking skills and the common absence of this productive skill as an essential component in school entrance exams. Moreover, in traditional English-speaking classes, classroom activities usually involve oral interactions between teachers and students, as well as among the students themselves. In addition, evaluating individual students' oral skills in a large class and providing instant feedback to each student may pose significant challenges for ESL teachers.

When assessing students' speaking proficiency, instructors may face the dilemma of whether to prioritize the evaluation of accuracy, complexity, or fluency in students' speaking abilities. For instance, Hulstijn (2011) differentiates between speaking accuracy as a component of basic linguistic proficiency, inherent in all native speakers, and complexity and fluency, which fall under the domain of extended linguistic proficiency achieved by advanced L1 and L2 learners. Furthermore, speaking proficiency is further influenced by the specific tasks assigned to second language learners (Blake, 2017). To provide better feedback to individual learners, Automatic Speech Recognition (ASR) has been proposed long ago. However, the accuracy of the machine-generated feedback has not been widely accepted by many students and teachers.

Currently, with the rapid development of machine learning and AI, the accuracy of several leading automated speech recognition (ASR) engines has significantly improved. These engines are widely available and play a transformative role in enhancing the speaking proficiency of L2 learners. One notable contribution is through personalized learning experiences, where AI analyzes individual learners' performance and provides clear, accurate feedback tailored to their specific needs. The integration of advanced ASR technology allows for real-time assessment of pronunciation, intonation, and fluency, offering learners immediate insights into their speaking skills. Interactive conversational agents, such as IPAs powered by AI, create authentic dialogue scenarios, enabling second language learners to engage in various conversations. Furthermore, ASR enables learners to participate in speaking exercises anytime, anywhere, breaking the constraints of traditional classrooms. This transformation extends beyond face-to-face conversation contexts and promotes L2 speaking in authentic and practical settings.

6.3 Automatic Speech Recognition System

Employing sophisticated algorithms and machine learning models, ASR (Automated Speech Recognition) systems are capable of analyzing and interpreting the inherent acoustic patterns of spoken words. This process involves deconstructing spoken language into phonemes and combining them with linguistic units to generate accurate transcriptions. The applications of ASR technology include intelligent personal assistants, voice-activated devices, transcription services, and language learning platforms. ASR holds the potential to transform human interactions with technology and information retrieval, enabling spoken language to become a crucial and easily accessible medium for human–computer communication. In L2 learning, ASR has been developed to provide valuable pronunciation feedback to students due to its ability to offer individualized guidance and immediate feedback—something that teachers may find challenging because of time constraints in the classroom (Ngo et al., 2024).

ASR (Automated Speech Recognition) specifically focuses on converting spoken words into written text. This technology supports individualized learning experiences by allowing learners to practice speaking, receive real-time assessments, and track their learning progress outside the classroom. However, researchers have reported varying findings on students' learning outcomes. For instance, McCrocklin (2019) found no significant difference in pronunciation performance between experimental and control groups. However, Gorjian et al. (2013) indicated that participants who practiced pronunciation with ASR significantly outperformed those who received traditional speaking training. Regarding effectiveness, Lai and Chen (2022) suggested that current ASR programs have become more accurate in transcribing non-native speech. Their findings also revealed that ASR programs were able to distinguish between proficient and less proficient learners, as low-proficiency learners had significantly lower accuracy rates compared to high-proficiency learners. In addition, Ngo et al. (2024) analyzed previous ASR studies and found that ASR is highly

Fig. 6.1 The speech assessment tool of Cool English site

effective for adult learners with an intermediate level of English proficiency. They also discovered that ASR-assisted pronunciation instruction in peer settings is more effective than using ASR independently. Regarding ASR feedback features, their findings indicate that pronunciation learning is more effective when ASR provides explicit corrective feedback, rather than simply transcribing words or offering correct/incorrect responses to speech input. Adult learners found implicit corrective feedback (e.g., model pronunciation with complete sentences or audio recast) more helpful, whereas younger learners, as reported by Wang and Young (2015), preferred explicit corrective feedback for self-correction.

Very recently, Microsoft launched the Speech Studio and a pronunciation assessment tool to evaluate learners' oral skills. The Cool English website (https://www.coolenglish.edu.tw) has closely followed this technology, releasing a series of complimentary English-speaking resources for ESL/EFL learners. The assessment tool provides instant, explicit feedback on four aspects of oral skills: pronunciation, accuracy, fluency, and completeness. Assessing speaking skills is often challenging for both teachers and students due to the time and effort required. However, this new ASR tool could make assessments more engaging and accessible to learners from diverse backgrounds. Figure 6.1 shows a screenshot of the Microsoft ASR technologies implemented on the Cool English website.

6.4 Speech Synthesis (TTS) for L2 Learning

Text To Speech (TTS) technology, closely linked to ASR, converts written text into speech with natural-sounding intonation and pronunciation. TTS technology shows great potential for enhancing L2 pronunciation teaching and learning, with research indicating that it is well-suited for L2 classrooms due to its human-like performance in intelligibility, comprehensibility, and morphophonemic accuracy (Cardoso et al.,

2015; Liakin et al., 2017). For example, TTS could provide auditory reinforcement to complement traditional reading activities, which offers significant benefits for L2 learners by enhancing their listening skills, phonetic awareness, and overall language comprehension. Furthermore, TTS exposes students to diverse accents and varieties of global English, which supports the goal of using English as a Lingua Franca in this global village. Beyond the classroom, the interactive nature of TTS enables students to practice pronunciation, rhythm, and intonation by mimicking the synthesized voices. This technology is considered valuable for self-paced learning, as it offers students the flexibility to listen to texts repeatedly until they grasp the intricacies of pronunciation and sentence structure.

Recently, the Cool English team has partnered with Microsoft's new Text-to-Speech (TTS) services to offer 36 different UK and US voice options for ESL/EFL users. As illustrated in Fig. 6.2, Cool English's TTS tool enables users to adjust the audio speed and download the recordings freely. With the rapid advancement of technology, the integration of ASR and TTS is also contributing to the development of Intelligent Personal Assistants (IPAs) and chatbots for L2 learning. The following section will explore the potential benefits and challenges of using IPAs in L2 education.

Fig. 6.2 TTS tool provided by Cool English

6.5 IPAs for L2 Learning

Intelligent Personal Assistants (IPAs) are voice-activated features integrated with smart speakers such as Google Assistant, Apple Siri, and Amazon Alexa. These services excel at performing concierge-like functions and delivering information, taking into account user input, location awareness, and the ability to retrieve data from a wide range of online sources (Tai & Chen, 2020). By utilizing ASR technology and NLP systems, IPAs interpret human language and respond to commands. According to data provided by Statista on the worldwide distribution of smart speaker market share from 2016 to 2022, the leading product currently is Amazon Echo, equipped with the IPA, Alexa. In the United States, more than half of consumers own a device with Alexa functionality or another digital smart assistant powered by AI (Bass, 2019). Alexa demonstrates versatility in performing a range of tasks, including internet searches, music playback, reading emails, connecting to social media, and even acting as a personal interpreter. Similarly, the Amazon store, like Google Play, offers a vast array of additional capabilities through the installation of extended commands (Yang et al., 2022).

According to Moussalli and Cardoso (2020), IPAs appear to be advantageous for language learning, as they potentially offer learners a substantial amount of input and enhance opportunities for oral interaction in the target language. Furthermore, IPAs enable learners to communicate in an environment with reduced anxiety, leading to increased engagement and willingness to communicate (Tai & Chen, 2020). Building on researchers' perspectives regarding IPAs, studies have explored the application of IPAs such as Alexa, Google Assistant, and Apple's Siri in supporting L2 learning. Empirical findings indicate that the use of IPAs enhance learners' oral skills (Dizon, 2020; Hsu et al., 2021; Yang et al., 2022), promote listening comprehension (Tai & Chen, 2022), and improve reading comprehension (Sing et al., 2019). Other findings have also revealed that IPAs reduce learners' speaking anxiety and promote students' willingness to communicate (Tai & Chen, 2020), as well as facilitate self-learning (Dizon & Tang, 2019). In the affective domain, Alexa has served as a helpful and non-threatening conversational partner for learners, reducing nervousness during interactions compared to face-to-face conversations with real people, and offering a private and comfortable setting (Sandeep, 2019). This learner-centered interaction and decreased inhibition may contribute to an improvement in communicative confidence.

While the affordances of IPAs have been verified by some empirical studies, their integration into ESL contexts still faces criticism. A notable issue is that IPAs often struggle to understand L2 learners' accented speech. Dizon (2017) revealed that the accuracy rate of Alexa's transcriptions of EFL learners' voice commands is only 50%. The effectiveness of IPAs in supporting ESL learners depends not only on the guidance provided by teachers but also on the learners' language proficiency. The less favorable outcomes in understanding the expressions of L2 learners could be linked to their limited proficiency in the target language (Wu et al., 2020). While numerous

studies have focused on the tertiary level, there remains a significant gap in understanding the impact of IPAs on younger learners. IPAs might provide a pleasant ubiquitous learning environment for ESL learners, but their role in supporting teachers in the classroom and facilitating students' learning outside the classroom warrants further investigation. As noted by Chen et al. (2023), a prerequisite for IPAs to adapt well to L2 learners' utterances is that the learners might need to reach a higher level of proficiency and pronunciation accuracy, as most ASR engines are not designed for non-native speakers (Levis, 2007). To apply IPAs effectively in ESL/EFL contexts, instructors need to invest in customized materials.

The findings indicate that while IPAs are effective in providing authentic input and facilitating simple interactions with L2 learners, their ability to engage in free-flowing conversations remains limited. IPAs have restricted knowledge bases and struggle to handle various topics raised by L2 learners. When L2 learners input ungrammatical sentences, IPAs may have significant difficulties in understanding their messages, which can lead to communication breakdowns and potential frustration for learners. Thus, IPAs are not yet capable of serving as fully competent interlocutors for L2 learners.

6.6 Chatbots in L2 Learning

A chatbot is an intelligent entity with the ability to converse with humans and interact with the surrounding world. More specifically, it is a computer program that leverages Natural Language Processing (NLP) technology, enabling the entity to engage in voice and text interactions with humans. The term "chatbot" originated from the Chatter Bot system, which was a computer-controlled player created to answer players' questions in the 1989 multiuser dungeon game "TinyMUD." This system was part of Mauldin's research in 1994.

Although "chatbot" was coined around 30 years ago, the concept of an intelligent system capable of engaging in conversation with human dates back to a much earlier period. In the 1960s, Joseph Weizenbaum from the Massachusetts Institute of Technology (MIT) developed ELIZA, the first computer program designed to engage in natural conversation with humans. The name "ELIZA" was derived from Eliza Doolittle, a character in George Bernard Shaw's "Pygmalion," as ELIZA's language abilities could be incrementally enhanced by its users. This paralleled Eliza Doolittle's transformation from a Cockney accent to an upper-class accent through expert language guidance. ELIZA simulated a psychotherapist interaction by transforming input text into interrogative forms, adopting pattern-matching techniques. It identified and processed keywords in the input and applied transformation rules; when no keywords were found, it modified the sentence by removing parts before a delimiter (Weizenbaum, 1966). Decades later, the Artificial Intelligence Markup Language (AIML), an XML-based language for building chatbots, was developed. A.L.I.C.E. (Artificial Linguistic Internet Computer Entity), created by Richard Wallace in 1995, was the first chatbot to utilize AIML. A.L.I.C.E. contained a "brain" made up of

41,000 categories, each linking a question (pattern) with a response (template), organized like branches on a tree. This structure allowed A.L.I.C.E. to efficiently match user questions with corresponding answers (Wallace, 2009).

Over time, the evolution of chatbots has taken significant strides, reaching a level that amazes even those living in this dynamic world. The growing interest in chatbots can largely be attributed to advancements in machine learning, as users increasingly explore their use for purposes beyond casual conversation, such as productivity, entertainment, and social interaction (Brandtzaeg & Følstad, 2017). Modern chatbots have found diverse applications across various domains, including education, information dissemination, customer service, and e-commerce. Some even assist with smoking cessation, like StopSmoking, or provide support for individuals experiencing stress and depression, such as Woebot (Croes & Antheunis, 2021).

In English language learning, socialbots and newly-developed language models might serve as 24/7 learning partners. For example, Mitsuku (now known as Kuki) (https://www.kuki.ai/), an AI-powered socialbot for English speakers, has interacted with approximately twenty-five million human end-users on online platforms and has won the Loebner Prize, a yearly Turing Test competition, five times. Kuki, developed mainly using Artificial Intelligence Markup Language (AIML), analyzes user input through a blend of statistical models, machine learning techniques, and manual tagging or review (ICONIQ, 2005). In addition, the chatbot also incorporates abuse detection and deflection strategies, ensuring smooth and respectful interactions. Users can engage with Kuki as if chatting with a real person, with the option to view its profile and learn more about it. Another notable socialbot is Chirpy Cardinal (https://stanfordnlp.github.io/chirpycardinal/), developed by Stanford University, which is designed to foster meaningful conversations. Considering aspects like knowledge, conversational style, and emotional connection, Chirpy Cardinal is designed to be cognitively and emotionally well-functioning. It can discuss a wide variety of topics with users and is also capable of responding empathetically to users' personal stories (Paranjape et al., 2020).

While these chatbots are beneficial for L2 learners, they often struggle to handle a wide range of topics introduced by users. For instance, Chirpy Cardinal can effectively respond to queries about the United States but encounters significant challenges when users ask questions related to Taiwan, which leads to inaccuracies and errors in brief interactions. This is illustrated in Fig. 6.3 below.

In November 2022, OpenAI launched ChatGPT, a large AI language model supported by Microsoft. The model quickly went viral since its debut, bringing the influence of artificial intelligence to individuals who had never used AI tools or were previously unfamiliar with the technology. ChatGPT's conversational style is versatile and designed to meet the diverse expectations of users. Unless specifically prompted, it typically delivers detailed responses to each prompt. In contrast to Mitsuku's short and sweet response, "Hi there! Quite a lot of things," ChatGPT gave a relatively comprehensive answer when asked, "Hi! What can we chat about?".

Hello! We can chat about a wide range of topics. Whether you're interested in discussing technology, science, literature, current events, or just looking for a friendly

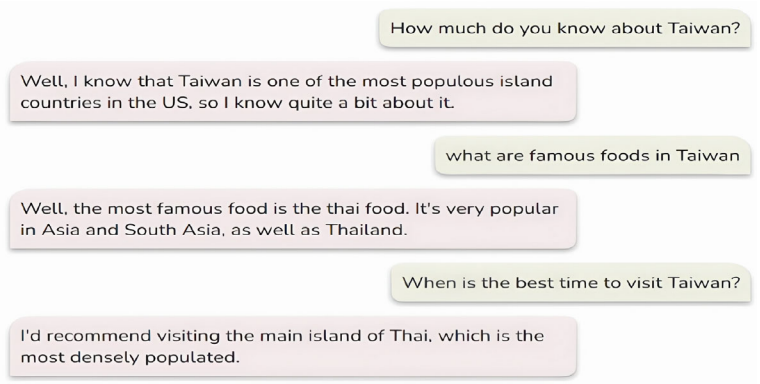


Fig. 6.3 The responses by Stanford Chirpy Cardinal Chatbot

conversation, I'm here for you. Is there a specific topic you have in mind, or do you have any questions you'd like to explore? (OpenAI, 2024).

When prompted appropriately, ChatGPT readily assumes various roles and engages with users from the perspective of those specific roles.

Chatbots exhibit great potential as a medium for language learning, serving both as tutors in practicing language and as independent learning tools (Haristiani, 2019). They offer contextual conversation reconstruction and help millions of learners in non-English-speaking settings to obtain opportunities to produce the target language (Tu, 2020). While the aforementioned bots are text-based and not specifically designed for language learning, they do not necessarily require users to "speak" and "listen." Text-based interactions with language partners have not been proven to facilitate users' speaking skills. To address this issue, CoolE Bot, a chatbot developed by National Taiwan Normal University (NTNU), was equipped with ASR and Speech Synthesis systems. By leveraging the Azure OpenAI Service and incorporating speech recognition and speech synthesis technologies, CoolE Bot enables users to engage in verbal interactions, allowing them to "talk" to the bot. The TTS and STT features distinguish CoolE Bot from other English socialbots like Mitsuku or Chirpy Cardinal, which indicates a stronger educational intention. To engage users, the developers have created chat rooms categorized by topics such as Cartoons, Stories, Jobs, and Games. In the chat room categorized as Cartoons, users can select which cartoon figures they wish to interact with. For instance, Fig. 6.4 illustrates how a cartoon figure, Joy, can engage with L2 students at their convenience.

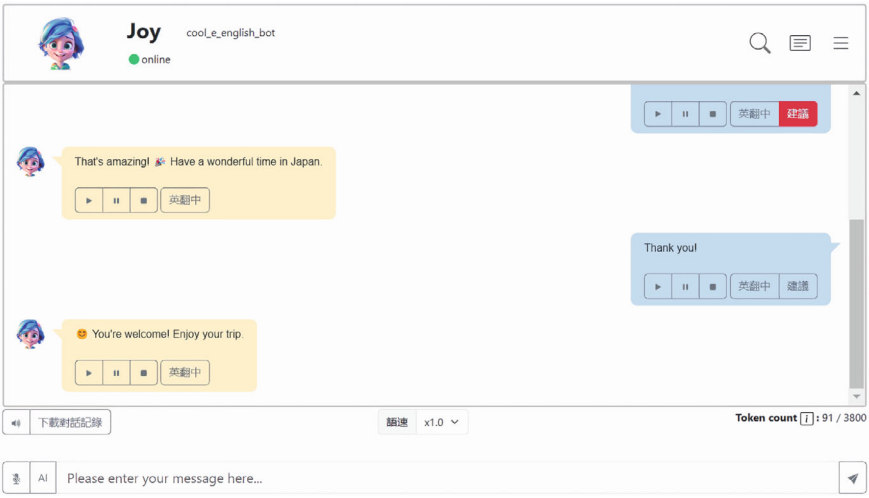


Fig. 6.4 Interacting with CoolEbot provided by Cool English

6.7 CALL/TELL and L2 Writing

In language learning, technologies designed for L2 writing have traditionally been categorized into three main groups: Web 2.0 applications, corpus-based tools, and Automated Writing Evaluation (AWE) systems (Li et al., 2017). Web 2.0 applications include social media platforms that enable multiple users to contribute content to a shared online space, such as Twitter and Facebook. These sites provide opportunities for users to engage with the language beyond the classroom (Li et al., 2017). Similarly, corpus-based tools provide users with access to authentic language use by enabling them to search a corpus for specific words or phrases. By contextualizing these words and phrases, a corpus helps users compare and contrast the usage of target vocabulary, clarifying how they are applied in different sentences. This is particularly useful for words with similar meanings that are challenging to explain through L1 translation alone, such as "high" and "tall." To assist users in producing more accurate and native-like writing, AWE tools have been developed. While these tools detect and correct mistakes or errors in texts, they have not yet achieved universal satisfaction, as not all users find them effective (Nunes et al., 2022). Developing a high-quality AWE system that performs like an actual English teacher, evaluating and correcting writing, remains a significant challenge. Moreover, the technology behind these systems largely depends on predicting common errors, while the potential for language errors is virtually limitless (Tschichold & Schulze, 2016).

Although developing a robust grammar checker has proven challenging, numerous researchers have identified contributions of these tools to L2 writing. For instance, automatic writing evaluation systems improve students' word usage in writing (Fitria, 2021; Ngo et al., 2022), foster self-directed learning (Lee, 2020) and enhance overall

writing quality (Zhai & Ma, 2023). However, the direct impact on users' writing skills per se remains uncertain, as indicated by a small effect on users' spelling, punctuation, and grammar (Ngo et al., 2022; Geng & Razali, 2022; Parra & Calero, 2019). Further empirical research is necessary in this area.

6.8 AI-Based Writing Tools

As previously discussed, creating a reliable grammar checker for any language poses significant challenges. L2 learners may make numerous types of errors, making it impossible to list all errors and train checkers to recognize each error pattern. Numerous grammar checkers are available, with Grammarly being one of the most highly recommended. This well-known writing assistant addresses various writing issues, including spelling, punctuation, and grammar. With advancements in NLP technologies, it now also provides guidance on content and organization. Studies have shown that Grammarly positively impacts L2 writing quality (Fitria, 2021) and is well-received by users (Pratama, 2021). Despite its fame and popularity compared to other grammar checkers like Ginger, Zoho Checker, and Quillbot, it has been indicated that Grammarly tends to provide more local feedback, focusing on vocabulary use and grammar, rather than global feedback on content and organizational aspects (Ngo et al., 2022).

While various automatic writing evaluation tools compete in the market, the emergence of ChatGPT appears to reshape the landscape of writing assistance. ChatGPT essentially functions as an online tool capable of performing numerous tasks for writers, such as answering questions, paraphrasing, summarizing, and generating tailored texts based on specific prompts. It not only performs tasks similar to traditional grammar checkers but often excels in a wider range of applications. The NTNU Cool English team has developed a robust grammar checker powered by ChatGPT which provides quick and detailed feedback on EFL learners' errors, offering sentence-by-sentence explanations for each mistake. An example of this functionality is illustrated in Fig. 6.5.

According to Chen (2023), ChatGPT's grammar-checking capabilities significantly surpass those of Grammarly in detecting common ESL errors. In a comparison involving over 500 errors from "Common Mistakes in English" (Fitikides, 2002) by Longman, ChatGPT 3.5 consistently outperformed the Grammarly Premium version. For example, Grammarly corrected only 22% of ungrammatical expressions, while ChatGPT successfully revised 75%. Furthermore, while Grammarly failed to correct any confused nouns, whereas ChatGPT achieved an impressive 86% success rate. For instance, it accurately transformed the phrase "Telling lies is a very bad custom" into "Telling lies is a very bad habit." Similarly, in addressing parts of speech confusion, Grammarly corrected just 43% of errors, while ChatGPT reached a 77% accuracy rate. In one case, Grammarly overlooked the correction of "You don't look as your mother," while ChatGPT revised it to "You don't look like your mother."

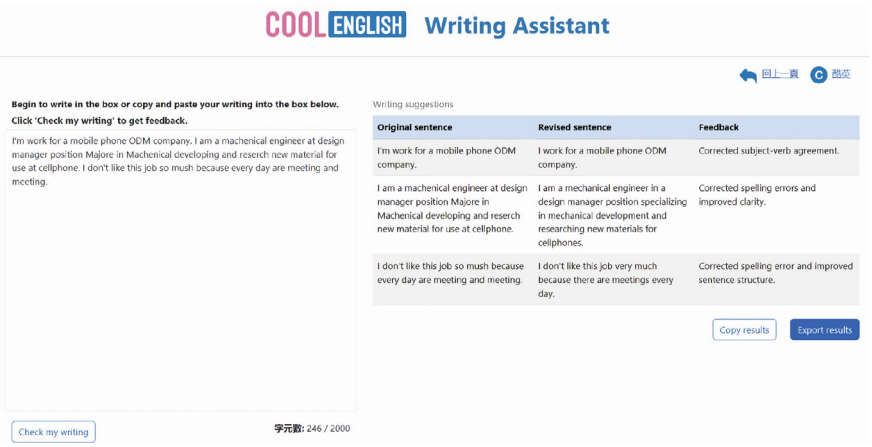


Fig. 6.5 The feedback and explanation provided at Cool English

In addition to its grammar-checking capabilities, ChatGPT's writing skills extend beyond surface-level tasks, delving into semantic, pragmatic, contextual, and stylistic issues in a writing piece and providing comprehensive feedback. What particularly draws users in is not only its quick response time but also the consistently high quality of its outputs. For instance, when asked to "Write a 200-word passage on the Importance of Sustainable Consumption," ChatGPT can produce a well-structured piece in under five seconds. The points raised include the reduction of carbon footprint and responsible resource management, which support the main idea and make it an exemplary piece on that topic. Users can also employ ChatGPT for more complex tasks, such as research writing and analyzing texts. ChatGPT's ability to interpret context beyond mere words enhances its effectiveness, which makes it a remarkable tool (Su et al., 2023).

6.9 Challenges in the AI Era

Artificial Intelligence is reshaping how individuals approach knowledge acquisition, providing access to a wide array of learning tools and resources via digital devices. However, the growing presence of misinformation and disinformation facilitated by AI reveals the need for users to adopt a more critical and discerning approach when engaging with such technologies. The ease of generating misleading content raises complex questions about the reliability of AI-driven information, which necessitates careful scrutiny of its use in educational contexts.

Beyond the risks posed by the spread of false and inappropriate content, learners face notable challenges in using AI tools for educational purposes. AI tools, such as ChatGPT, may inadvertently misattribute concepts (Dergaa et al., 2023) or generate

seemingly realistic sensory experiences that do not correspond to any real-world input (Alkaissi & McFarlane, 2023). These issues stem from the limitations in the AI's training data, transformer architecture, and reinforced learning from human feedback (Hwang et al., 2023). While AI's convenience offers clear advantages, excessive reliance on these technologies may potentially hinder the development of learners' independent skills. On the other hand, AI and NLP technologies, including ChatGPT, are not immune to bias, which may influence the quality and accuracy of the text they generate. In language education, such biases may inadvertently affect students' learning and development (Baskara, 2023). These concerns also raise ethical questions: Should large language models like ChatGPT be credited as co-authors in academic writing, or does their use for paraphrasing constitute a form of plagiarism?

Similarly, AI tools also present challenges for teachers. Integrating AI technologies to support language learning while mitigating their potential misuse is a complex task. Teachers at all levels may require additional training and support to effectively incorporate AI into their educational practices. While the development of AI plagiarism detectors is advancing alongside AI-assisted writing tools, addressing these challenges can feel like trying to extinguish a fire without knowing how close the water source is. Reaching a consensus and establishing a strategic approach are crucial to navigating the plagiarism concerns that accompany these advanced tools. In this context, teachers' guidance becomes essential in helping students use AI responsibly for language learning.

The transition from the early stages of CALL/TELL to the era of AI-assisted language learning marks a significant shift in the research agenda. AI integration has reshaped second and foreign language learning by enabling personalized, adaptive, and technology-enhanced learning experiences. This restructuring has opened up opportunities for more tailored approaches to language acquisition, driven by the dynamic capabilities of AI. As advancements continue, innovative AI-based CALL/TELL tools are expected to further influence, transform language education, and expand opportunities for individualized learning.

Acknowledgements The research was supported by the National Science and Technology Council, Taiwan, under project numbers 110-2410-H-003-019-MY3 and MOST 113-2410-H-003-071-MY2.

References

Alkaissi, H., & McFarlane, S. I. (2023, February 19). Artificial hallucinations in ChatGPT: Implications in scientific writing. *Cureus, 15*(2).

Baskara, F. R. (2023). Integrating ChatGPT into EFL writing instruction: Benefits and challenges. *International Journal of Education and Learning, 5*(1), 44–55.

Bass, L. (2019). The concealed cost of convenience: Protecting personal data privacy in the age of Alexa. *Fordham Intell. Prop. Media & Ent. LJ, 30*, 261.

Blake, R. J. (2017). Technologies for teaching and learning L2 speaking. *The handbook of technology and second language teaching and learning*, 107–117.

Brandtzaeg, P. B., & Følstad, A. (2017). Why people use Chatbots. In *Internet Science: 4th International Conference, INSCI 2017, Thessaloniki, Greece, November 22–24, 2017, Proceedings 4* (pp. 377–392). Springer International Publishing.

Cardoso, W., Smith, G., & Garcia Fuentes, C. (2015, December). Evaluating text-to-speech synthesizers. In *Critical CALL–Proceedings of the 2015 EUROCALL Conference, Padova, Italy* (pp. 108–113). Research-publishing. net

Chen, H. (2023) *Comparing the Performances of Grammarly and ChatGPT in Correcting ESL Writing Errors.* 2023 presentation at the 15th Asian Conference on Education (ACE2023), Tokyo.

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

Croes, E. A., & Antheunis, M. L. (2021). Can we be friends with Mitsuku? A longitudinal study on the process of relationship formation between humans and a social chatbot. *Journal of Social and Personal Relationships, 38*(1), 279–300.

Dergaa, I., Chamari, K., Zmijewski, P., & Saad, H. B. (2023). From human writing to artificial intelligence generated text: Examining the prospects and potential threats of ChatGPT in academic writing. *Biology of Sport, 40*(2), 615–622.

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

Dizon, G., & Tang, D. (2019). A pilot study of Alexa for autonomous second language learning. *CALL and complexity*, 107.

Dizon, G. (2020). Evaluating intelligent personal assistants for L2 listening and speaking development.

Dizon, G., & Gayed, J. M. (2021). Examining the impact of Grammarly on the quality of mobile l2 writing. *Jalt Call Journal*, *17*(2), 74–92.

Fitikides, T. J. (2002). *Common mistakes in English: With exercises* (6th ed.). Longman.

Fitria, T. N. (2021). Grammarly as AI-powered English writing assistant: Students' alternative for writing English. *Metathesis: Journal of English Language, Literature, and Teaching, 5*(1), 65–78.

García, C., Nickolai, D., & Jones, L. (2020). Traditional versus ASR-based pronunciation instruction. *Calico Journal, 37*(3), 213–232.

Geng, J., & Razali, A. B. (2022). Effectiveness of the automated writing evaluation program on improving undergraduates' writing performance. *English Language Teaching, 15*(7), 49–60.

Gorjian, B., Hayati, A., & Pourkhoni, P. (2013). Using Praat software in teaching prosodic features to EFL learners. *Procedia-Social and Behavioral Sciences, 84*, 34–40.

Haristiani, N. (2019, November). Artificial intelligence (AI) chatbot as language learning medium: An inquiry. In *Journal of Physics: Conference Series* (Vol. 1387, No. 1, p. 012020). IOP Publishing.

Hulstijn, J. H. (2011). Language proficiency in native and nonnative speakers: An agenda for research and suggestions for second-language assessment. *Language Assessment Quarterly, 8*(3), 229–249.

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.

Hsu, H. L., Chen, H. H. J., & Todd, A. G. (2021). Investigating the impact of the Amazon Alexa on the development of L2 listening and speaking skills. *Interactive Learning Environments, 31*(9), 5732–5745.

Hwang, G. J., Xie, H., Wah, B. W., & Gašević, D. (2020). Vision, challenges, roles and research issues of Artificial Intelligence in Education. *Computers and Education: Artificial Intelligence, 1*, Article 100001.

Hwang, S. I., Lim, J. S., Lee, R. W., Matsui, Y., Iguchi, T., Hiraki, T., & Ahn, H. (2023). Is ChatGPT a "Fire of Prometheus" for non-native English-speaking researchers in academic writing? *Korean Journal of Radiology, 24*(10), 952.

ICONIQ. (2005). *Iconiq*. https://www.iconiq.ai/

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

Lai, C., & Gu, M. (2011). Self-regulated out-of-class language learning with technology. *Computer Assisted Language Learning, 24*(4), 317–335.

Lai, K. W. K., & Chen, H. J. H. (2022). An exploratory study on the accuracy of three speech recognition software programs for young Taiwanese EFL learners. *Interactive Learning Environments*, 1–15.

Levis, J. (2007). Computer technology in teaching and researching pronunciation. *Annual Review of Applied Linguistics, 27*, 184–202.

Li, Z., Dursun, A., & Hegelheimer, V. (2017). Technology and L2 writing. *The handbook of technology and second language teaching and learning*, 77–92.

Liakin, D., Cardoso, W., & Liakina, N. (2017). The pedagogical use of mobile speech synthesis (TTS): Focus on French liaison. *Computer Assisted Language Learning, 30*(3–4), 325–342.

Lee, Y. J. (2020). The long-term effect of automated writing evaluation feedback on writing development. *English Teaching, 75*(1), 67–92.

Lin, M.P.-C., & Chang, D. (2020). Enhancing post-secondary writers' writing skills with a chatbot. *Journal of Educational Technology & Society, 23*(1), 78–92.

McCrocklin, S. (2019). ASR-based dictation practice for second language pronunciation improvement. *Journal of Second Language Pronunciation, 5*(1), 98–118.

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

Ngo, T. T. N., Chen, H. H. J., & Lai, K. K. W. (2022). The effectiveness of automated writing evaluation in EFL/ESL writing: A three-level meta-analysis. *Interactive Learning Environments*, 1–18.

Ngo, T. T. N., Chen, H. H. J., & Lai, K. K. W. (2024). The effectiveness of automatic speech recognition in ESL/EFL pronunciation: A meta-analysis. *ReCALL, 36*(1), 4–21.

Nunes, A., Cordeiro, C., Limpo, T., & Castro, S. L. (2022). Effectiveness of automated writing evaluation systems in school settings: A systematic review of studies from 2000 to 2020. *Journal of Computer Assisted Learning, 38*(2), 599–620.

OpenAI. (2024). ChatGPT (January 16 version) [Large language model]. https://chat.openai.com/chat

Paranjape, A., See, A., Kenealy, K., Li, H., Hardy, A., Qi, P., Sadagopan, K. R., Phu, N. M., Soylu, D., & Manning, C. D. (2020). Neural generation meets real people: Towards emotionally engaging mixed-initiative conversations. *arXiv preprint arXiv:2008.12348.*

Parra, G., & L., & Calero S, X. (2019). Automated writing evaluation tools in the improvement of the writing skill. *International Journal of Instruction, 12*(2), 209–226.

Pratama, Y. D. (2021). The investigation of using Grammarly as online grammar checker in the process of writing. *English Ideas: Journal of English Language Education, 1*(2).

Richards, J. C. (2015). The changing face of language learning: Learning beyond the classroom. *Relc Journal, 46*(1), 5–22.

Sandeep, S. S. (2019). Learning English language through Amazon Alexa for Indian students. *Journal of the Gujarat Research Society, 21*(10), 619–622.

Sing, P. B., Embi, M. A., & Hashim, H. (2019). Ask the Assistant: Using Google assistant in classroom reading comprehension activities. *Int. J. New Technol. Res, 5*(7), 39.

Su, Y., Lin, Y., & Lai, C. (2023). Collaborating with ChatGPT in argumentative writing classrooms. *Assessing Writing, 57*, Article 100752.

Tai, T. Y., & Chen, H. H. J. (2022). The impact of intelligent personal assistants on adolescent EFL learners' listening comprehension. *Computer Assisted Language Learning*, 1–28.

Tai, T. Y., & Chen, H. H. J. (2020). The impact of Google assistant on adolescent EFL learners' willingness to communicate. *Interactive Learning Environments, 31*(3), 1485–1502.
Tschichold, C., & Schulze, M. (2016). Intelligent CALL and written language. In F. Farr & L. Murray (Eds.), *The Routledge handbook of language* (pp. 522–535). Taylor & Francis.
Tu, J. (2020). Learn to speak like a native: AI-powered Chatbot simulating natural conversation for language tutoring. *Journal of Physics: Conference Series, 1693*(1), Article 012216. https://doi.org/10.1088/1742-6596/1693/1/012216
Wallace, R. S. (2009). *The anatomy of ALICE* (pp. 181–210). Springer.
Wang, Y. H., & Young, S. C. (2015). Effectiveness of feedback for enhancing English pronunciation in an ASR-based CALL system. *Journal of Computer Assisted Learning, 31*(6), 493–504.
Wang, Y. F., Petrina, S., & Feng, F. (2017). VILLAGE—Virtual immersive language learning and gaming environment: Immersion and presence. *British Journal of Educational Technology, 48*(2), 431–450.
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.
Wu, Y., Edwards, J., Cooney, O., Bleakley, A., Doyle, P. R., Clark, L., Rough, D., & Cowan, B. R. (2020, July). Mental workload and language production in non-native speaker IPA interaction. In *Proceedings of the 2nd Conference on Conversational User Interfaces*, 1–8.
Yang, C. T. Y., Lai, S. L., & Chen, H. H. J. (2022). The impact of intelligent personal assistants on learners' autonomous learning of second language listening and speaking. *Interactive Learning Environments*, 1–21.
Zhai, N., & Ma, X. (2023). The effectiveness of automated writing evaluation on writing quality: A meta-analysis. *Journal of Educational Computing Research, 61*(4), 875–900.
Zhao, Y., & Lai, C. (2023). Technology and second language learning: Promises and problems. In *Technology-mediated learning environments for young English learners* (pp. 167–206). Routledge.

Howard Hao-Jan Chen is a distinguished professor of the English Department at National Taiwan Normal University, Taipei, Taiwan. He also serves as the associate editor of the ETS (Educational Technology and Society) journal. In addition, he is also on the editorial board of Computer-Assisted Language Learning Journal. Professor Chen has published many papers in Computer-Assisted Language Learning, ReCALL, and several related language learning journals. His research interests include computer-assisted language learning, corpus research, and second language acquisition. He also manages a large national English learning website called Cool English (www.coolenglish.edu.tw). The website supported by Ministry of Education has more than 2.1 million registered users.

Mr. Hsieh Fu-Chan has graduated from the TESOL program of National Taiwan Normal University and they are now teaching in senior high schools in Taiwan.

Miss Yu-Ci Yen has graduated from the TESOL program of National Taiwan Normal University and they are now teaching in senior high schools in Taiwan.

Chapter 7
Emulating Real World Environment for Contextual Language Learning Using Educational Robot and IOT-Enabled Tangible Objects

Nian-Shing Chen and Hui-Chin Yeh

Abstract This chapter introduces a pioneering approach to foreign language acquisition using an Emulated Real-World Environment, combining Educational Robots and IoT-Enabled Tangible Objects. Addressing critical challenges in language learning, such as lack of immersion and contextual practice, it introduces the R&T (Robot and Tangible objects) learning system. This system integrates humanoid robots with IoT-enabled tangible objects, facilitating realistic, context-rich language interactions. It supports diverse learning modes and focuses on enhancing conversational skills, digital literacy, and physical concept comprehension. The research highlights an interdisciplinary collaboration that ensures educational effectiveness, technological sophistication, and contextual relevance. A key outcome is significant improvement in language comprehension and communication skills, demonstrated through multi-sensory learner interactions. Additionally, the study explores applications in early childhood language development, using AI-driven educational robots to address language gaps. This innovative approach marks a transformative step in language education, merging robotics and AI to enrich language-learning experiences.

Keywords Educational robot · Tangible object · Physical teaching aid · Language learning · Contextual learning environment

N.-S. Chen (✉)
Institute for Research Excellent in Learning Sciences, Program of Learning Sciences, National Taiwan Normal University, Taipei City, Taiwan
e-mail: nianshing@gmail.com

H.-C. Yeh
Department of Applied Foreign Languages, National Yunlin University of Science and Technology, Douliu, Taiwan
e-mail: hyeh@yuntech.edu.tw

Y. Lan et al. (eds.), *AI-Mediated Language Education in the Metaverse Era*, Chinese Language Learning Sciences, https://doi.org/10.1007/978-981-95-0245-5_7

7.1 Introduction

7.1.1 The Challenges of Learning a Foreign Language

The acquisition of a foreign language presents several challenges that hinder learners' progress. One primary challenge is the absence of an appropriate language-learning environment (Renandya, 2013). Many individuals opt to immerse themselves in a foreign country to facilitate more effective language acquisition (Tseng et al., 2021). However, simply residing in a foreign environment is insufficient, as learners may still struggle due to the lack of authentic contexts in which to apply their language skills (He & Deng, 2015). Without exposure to diverse real-life situations, learners miss critical opportunities to practice and reinforce their linguistic competence in meaningful ways (Fryer et al., 2006; Renandya, 2013). A third key challenge relates to the learner's ability or flexibility in engaging in oral communication (Horwitz et al., 1986). Even when placed in an appropriate contextual environment, learners who do not actively participate in conversations experience restricted language practice. Consequently, the absence of interaction within the target language environment significantly impedes linguistic development (Fryer et al., 2006).

7.1.2 The Solution Concepts for Tackling Foreign Language-Learning Challenges

To address the challenge of a lacking suitable environment, the concept of an emulated real-world environment is proposed (Chen, 2016; Milton et al., 2012). Since not all learners can travel abroad, technology-driven simulations tailored to specific topics or professional domains provide an immersive learning experience (Zhao, 2021). For example, replicating an agricultural farm or zoo using physical models, such as LEGO blocks and toy objects, allows learners to engage in interactive, context-rich learning.

To address the lack of meaningful contextual exposure, structured scenario-based learning can be integrated into these simulated environments. Real-world tasks, such as managing soil, water, and nutrients in agriculture or feeding animals in a zoo, require both linguistic and domain-specific knowledge, reinforcing language skills through practical application (Wahl & Winiwarter, 2012). These situational learning contexts enhance comprehension and retention by bridging the gap between theoretical knowledge and real-world use (Tutal, 2023).

To solve the challenge of limited capacity or flexibility, educational robots can offer one-on-one interactions (Konijn et al., 2022), facilitating on-demand and self-paced learning in these emulated real-world environments. This flexibility allows learners to take full ownership of their learning, enhancing their agency and autonomy (Hsieh et al., 2023; Lin et al., 2022; Peura et al., 2023), but also sets a foundation for incorporating other anxiety-mitigating educational technologies. In this context, the

findings from Yang et al. (2024) are particularly relevant, demonstrating how digital game-playing can significantly enhance English vocabulary and content learning among anxious learners, thus complementing our R&T system's capabilities by broadening the ways in which technology can support diverse learner needs (Denden et al., 2024).

7.2 The Robot and Tangible Object (R&T) Learning System

The robot and tangible object (R&T) learning system integrates (a) a humanoid educational robot to facilitate social and verbal exchanges, (b) tangible objects such as IoT-enabled physical aids that support hands-on experiences, and (c) a digital interface (e.g., a tablet) to display text prompts, multimedia, and transcripts (Lin et al., 2021). By offering verbal, digital, and tangible interaction, the system supports embodied cognition and multimodal language practice.

7.2.1 Interaction Modalities in an R&T Learning System

In an R&T learning system, three primary interaction modalities are structured, each targeting distinct cognitive and linguistic processes:

1. **Learner-Robot Interaction:**
   In this modality, the robot serves as a conversational partner, prompting learners to practice dialogue in a supportive environment. Research shows that conversing with social robots can reduce learners' anxiety and enhance retention (Kory Westlund et al., 2017; Lin et al., 2022).
2. **Learner-Digital Content (Tablet) Interaction:**
   The learner employs a tablet or similar device to view multimedia learning materials and receive immediate feedback on accuracy (Mayer, 2005; Sweller, 2011). Animated visuals, subtitles, or quizzes on the tablet further reinforce vocabulary, grammar, and reading comprehension (Hu et al., 2023).
3. **Learner-Physical Teachin Aid (IoT Tangible Object) Interaction:**
   Tangible items, such as NFC-tagged blocks or flashcards, enable kinesthetic engagement. Learners touch or scan objects following robot instructions, which deepens semantic understanding by linking language concepts to real-world referents (Hsu et al., 2022; Hu et al., 2020).

A tangible example of integrating the three interaction modalities within the R&T system is vocabulary learning through a toy teddy bear model. This application follows a structured, multi-step approach:

- Learner-Robot Interaction: The robot introduces the activity and prompts the learner to touch sensors on its body parts, beginning with the head. Upon interaction, the robot states, "This is my head. Can you say, 'This is my head'?" The learner's response is analyzed using Speech-to-Text technology, and the robot provides real-time corrective feedback.
- Learner-Digital Content Interaction: After the verbal exchange, the robot presents a multiple-choice question on the tablet, prompting the learner to identify the word "head." This visual reinforcement strengthens word recognition.
- Learner-Tangible Object Interaction: The robot then instructs the learner to scan an NFC tag embedded in the teddy bear's head. This step ensures that the learner associates the vocabulary word with its real-world referent, reinforcing symbolic-to-concrete language learning.

By integrating verbal, digital, and tactile interactions, this learning sequence deepens comprehension, supports multimodal engagement, and enhances retention.

7.3 Implementation Modes and Learning Application Design of R&T Learning System

7.3.1 Three Learning Modes

1. Physical Mode: In this mode, learners are in a physical classroom where they interact directly with the R&T learning system. This mode is typical for traditional classroom settings, where students engage with the physical components of the system.
2. Online Mode: Learners participate from different locations through an online classroom. In this setup, the R&T learning system is located at the teacher's end. Students use an R&T Remote App to interact with the robot, digital content, and tangible objects.
3. Blending Mode: This is a hybrid approach where some students are present in the physical classroom while others participate online. For instance, students who are unable to attend school due to illness like COVID-19 can join the session online from home, ensuring continuity in their learning experience.

7.3.2 Designing R&T Learning Application

The structure of an R&T learning application is systematically organized to optimize learning efficiency. Here's a breakdown:

1. Learning Activity (LA): Each LA is centered around a specific topic with a clear learning goal, designed to be completed within a reasonable timeframe based on the learner's age group.
2. Learning Parts (LP): Each LP within an LA serves a distinct purpose. For example, Part I might introduce the activity, Part II could involve guided learning for basic knowledge acquisition, Part III might focus on self-directed learning to practice or apply the knowledge, and Part IV could be for assessment and feedback.
3. Learning Tasks (LT): Tasks within each LP involve interactions between the learner and the robot. These could be tasks initiated by the robot, such as asking a question or instructing the learner to sort objects by color, or they could be initiated by the learner to engage with the robot.
4. Physical Actions (PA): PAs are actions executed either by the robot or the learner to complete a task. For instance, in a task where the robot asks a question, the sequence of actions includes the robot posing the question, the learner responding, and the robot providing feedback based on the correctness of the answer.

This structure ensures a comprehensive and engaging learning experience, incorporating both interactive and reflective elements.

7.3.3 Procedure for R&T Application Development

Designing and implementing an R&T learning application involves close collaboration among three teams:

1. Content Domain Expert Team (CK): Create and refine the conceptual design and learning objectives.
2. Pedagogy Expert Team (PK): Translate the content design into an R&T script that is instructionally sound.
3. Technology Expert Team (TK): Code the final application (APK) based on the R&T script, ensuring proper functionality.

These three interdisciplinary stakeholders and their tasks are interconnected, each playing a vital role in developing an engaging R&T learning application (Table 7.1).

Table 7.1 The three interdisciplinary stakeholders and their tasks

Stakeholders	Tasks
Content experts	Responsible for creating the R&T conceptual design document of the learning activity, focusing on the subject matter and core content
Pedagogy experts	Tasked with developing the R&T script design document based on the conceptual design. Their role is to ensure that the educational strategy is effectively integrated
Technology experts	In charge of coding the R&T learning application (APK), translating the R&T script into a functional technology product

7.4 Design Principles and Guidelines for Creating R&T Learning Applications

7.4.1 R&T Conceptual Design Principles

1. Balanced media usage: The use of robot, tablet, and physical tangible object in a learning activity should be balanced, avoiding overemphasis on one type of media (Hwang & Chen, 2023).
2. Structural learning activities: The structure of a learning activity should include various parts, each comprising several tasks, which should contain specific actions (Eseryel et al., 2011).
3. Action-specific objectives: An action could involve interacting with a robot, a tablet, or a tangible object. The choice of interaction type should be based on the learning objective of that particular action (Ifenthaler et al., 2011).
4. Clear version control: File names should follow this format "Script-ScriptName-YYYYMMDD-AuthorName-Version-latest." For example, "Script-EmotionDetective-20240120-js-v1-latest." All versions should be stored in a cloud folder, with older versions archived in an "old" folder.

7.4.2 R&T Conceptual Design Guidelines

1. Robot acknowledgment: The robot should always acknowledge or respond to the learner's actions/inputs to maintain conversational coherence.
2. Robot output attributes: Every robot action includes (a) speech content, (b) hand gestures, (c) facial expressions, (d) speech pace, and optionally (e) music or sound effects. If unspecified, default settings apply to all except speech content.
3. Standby screens on tablets: Displaying a relevant, visually appealing screen when no explicit content is shown helps maintain thematic immersion.
4. Detailed NFC tag definitions: When defining an NFC tag, it's best to include all metadata of a physical object, using "/" as a separator. For example, an apple's NFC tag could be: apple/蘋果/red/round/vitamin C.

5. Variable use in dialogue: "If a variable (a content with multiple possible values) is used in a dialogue, a list (…) should be used, with values separated by." For example, to describe three students with class, seat number, and name: (5/13/Guan-Bo, 1/2/Xiao-Ming, 3/4/Da-Ming).
6. Learner input scenarios: When awaiting learner input after a task or question, the following scenarios should be considered:

    a. Correct answer-Proceed accordingly.
    b. Wrong answer-Provide corrective feedback or hints.
    c. Not a valid input—If the response is outside predefined options (e.g., instructed to touch left or right hand, but another part is touched), prompt clarification or guidance.
    d. No response and time out—If not input is received within a set time, issue a reminder or rephrase the prompt.

For example, for understanding game rules, responses should include:

- Yes (understand)—Proceed with the game.
- No (don't understand)—Robot provides a more detailed explanation.
- Invalid response—Robot re-explains the rules with clarification.
- No response within time limit—Robot issues a reminder.
- Session closure: At the end of the learning activity, robot should always offer the learner to play again or end the learning session.

7.4.3 R&T Script Design Principles

1. Clarity and specificity: The script should be clearly written, detailing the learner's speech and actions, followed by the educational robot's assessment and response.
2. Versional file naming: Every part, task, and action described in the R&T conceptual design document must correspond to the script.
3. Compliance with computational thinking: File naming format: "Script-ScriptName-Date-Author-Version-latest." (Mark "latest" only for the most recent version). For example, this script should be named: "Script-EmotionDetective-20240120-kp-v1-latest."
4. Robot feature configuration: All versions should be retained in a cloud folder, with older versions stored in a separate "old" folder.

7.4.4 R&T Script Design Guidelines

1. Use of lowercase: Scripts should be written in lowercase letters unless uppercase is specifically required and marked.
2. Computational thinking skill: Script designers should have a basic understanding about computational skill for script writing. For repetitive tasks, write a logical

procedure using variables. Example: Let X be a specific meal from five different options of A, B, C, D, E in which a student could choose from X.
3. Traceable revisions: When making any changes to the script document, making a clear tracking marks and/or comment so that coders could find those changes easily. Remove those markers once the implemented APK and the script design contents are synchronized.
4. Script accuracy: The robot's features must be correctly set to generate accurate responses, such as enabling action functionality.
5. Script end option: Offer learners a choice to play again or exit before end the activity.
6. Script start trigger: A clear physical trigger (e.g., touching the robot's right cheek) initiates the application.

7.4.5 R&T Script Design Checklist

1. Logical consistency: Whether the script content has been mentally rehearsed step by step to ensure logical consistency.
2. Handling of unexpected inputs: Whether some potential issues from unexpected user's inputs and prevention mechanism have been considered.
3. Avoidance of redundancies: Whether the script contains any meaningless or unusual actions or steps.
4. Cross-checking: Whether cross-checking have been carried out among script designers for logical consistency before handing it over to coders.

7.5 An Illustrative R&T System Example

One example of an R&T implementation appears in Hsieh et al. (2023), where a system combines a humanoid robot called Kebbi, a tablet, a smartphone, and tangible objects (laminated cards). Kebbi's expressive faces and gestures form the emotional and conversational core, while scanning the physical cards provides tactile reinforcement. The tablet displays learning content (e.g., transcripts of Kebbi's speech), creating a complete multimodal context (Cheng et al., 2022; Chiang et al., 2023).

7.5.1 System Architecture

Figure 7.1 shows the system architecture, which uses a Message Queuing Telemetry Transport (MQTT) and Internet of Things (IoT) Server, learning applications, and a physical robot named Kebbi, accompanied by a tablet and tangible objects. The MQTT and IoT Server, using Home Assistant, serves as a communication hub linking

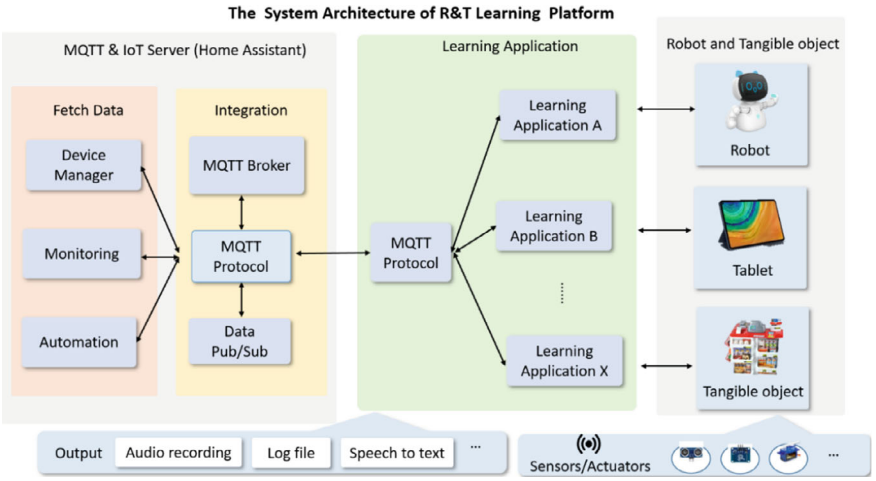


Fig. 7.1 The system architecture of R&T learning system

robot sensors, the tablet interface, and the learners' devices. Kebbi, equipped with sensors and designed to exhibit context-appropriate facial expressions and movements, interacts with students to facilitate learning. The tangible objects, such as laminated cards with English words and corresponding pictures, are scanned using a cellphone, further engaging students in the learning process.

7.5.2 Main Components and Interaction Examples

The R&T system consists of a physical robot (Kebbi), a tablet, a mobile phone and tangible objects as shown in Fig. 7.2. In the R&T system, Kebbi is used as student's interlocutor that also shows context-appropriate facial expressions and physical movements while the tablet is being used as a medium to display multimodal learning contents, including Kebbi's spoken sentences, pictures, speech-to-text ASR results, and conversation hints. The tangible objects in the system are laminated cards with a picture and its corresponding English word at the front, and QR code at the back. In order for the students to interact with tangible objects in the system, a cellphone is used as QR code scanner to scan the QR codes on the back of the tangible objects (cards).

Kebbi is equipped with sensors on different parts of its body. Figure 7.3-L shows the touchable sensors on Kebbi's hands along with the tablet displaying multimodal learning contents to help with comprehension of Kebbi's sentences, while Fig. 7.3-R illustrates Kebbi's gesture and expression when Kebbi talks.

The integration of these components creates a highly interactive and immersive learning environment, where robotics and IoT-based tangible objects work together

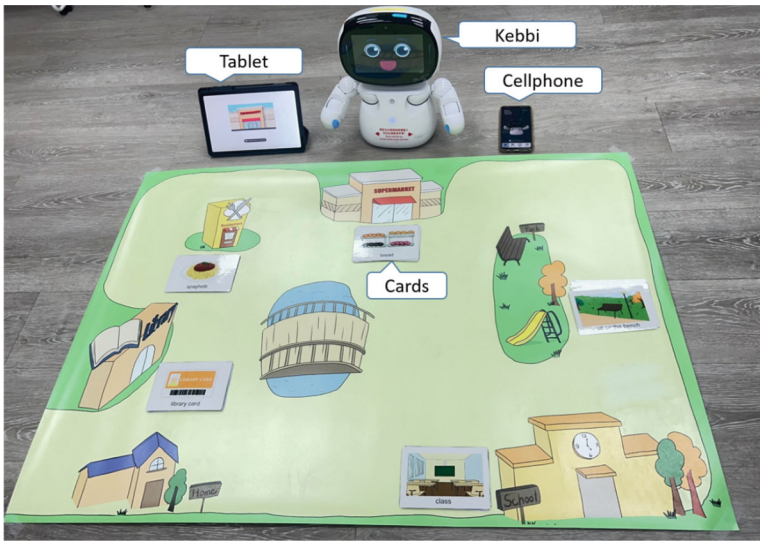


Fig. 7.2 The main components of an R&T system

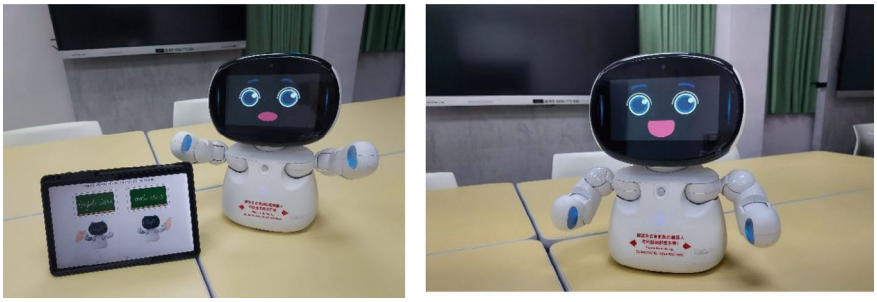


Fig. 7.3 L: The touch sensors on Kebbi. R: Kebbi's gesture and expression

to enhance language acquisition. Previous research has demonstrated the effectiveness of this system in facilitating contextual learning and improving learner engagement. For example, Cheng et al. (2022) investigated an English vocabulary game in which learners interacted with Kebbi and IoT-based tangible objects. Their study found that Kebbi's speech, gestures, and facial expressions, combined with QR-coded laminated cards, helped learners reinforce vocabulary acquisition. The physical interaction with the robot's touch-sensitive sensors and the scanning of tangible objects allowed students to engage with the learning material in a more meaningful way. Similarly, Chiang et al. (2023) examined students' observable behaviors while interacting with Kebbi and tangible objects. Their analysis identified instances where learners relied on Kebbi's multimodal cues (gestures, expressions, and speech) to understand instructions and complete language tasks. The

study highlighted how learners engaged with the system by combining physical touch with digital input, demonstrating the effectiveness of robot-tangible object interaction in creating an engaging and context-rich learning experience. These studies illustrate how the seamless integration of robots and IoT-enabled tangible objects fosters active learner engagement and contextualized language learning, reinforcing the effectiveness of the R&T system.

7.5.3 Learning Activity Design

The learning activity design is centered around creating a natural, conversational environment for language practice, particularly in pronunciation. The R&T system is tailored to mimic real-world conversation dynamics, where the robot acts as a patient and flexible interlocutor, offering learners repeated opportunities to practice speaking. The learning activities are set in a virtual "Happy Town," with tasks designed around everyday situations in various locations like a supermarket, library, or park. These tasks are based on the Task-Based Learning (TBL) approach, emphasizing meaningful, communicative use of language. (Fang et al., 2021; Wenas et al., 2023) The system aims to increase learners' confidence and willingness to communicate by allowing them to interact with the robot and use tangible objects to facilitate understanding and pronunciation of target words and phrases.

The study meticulously evaluates the R&T system's effectiveness through pre- and post-tests, alongside video analyses of student interactions, providing a comprehensive insight into the system's impact on language learning. Notable improvements in students' English pronunciation and willingness to communicate highlight the system's efficacy, although challenges in robot responsiveness were noted, offering directions for future enhancements. Building on this, Cheng et al. (2022) further illuminate the potential of such technologies, demonstrating how the integration of robots and IoT-based tangible objects significantly enhances game-based language learning for children. This dual evidence underscores the effectiveness of interactive tools in fostering not only engagement but also linguistic comprehension, thereby affirming the utility of our R&T learning system across diverse educational settings.

7.6 Building a Smart Storytelling Robot by Integrating Multimodal Generative AI Technologies

A separate but related line of research involves integrating generative AI (e.g., ChatGPT) into storytelling robots for children ages 3–7 (Lin et al., 2022). Such systems aim to automate personalized story creation and synchronize robot gestures, facial expressions, and supplementary multimedia content, thereby creating a more

context-rich and engaging narrative experience. By leveraging ChatGPT's application programming interface, the robot can generate customized stories aligned with individual learners' interests, while interactive quizzes and games further foster critical thinking, problem-solving, and memory retention. This approach also addresses the "thirty-million-word gap," promotes a sustained interest in learning, and broadens access to high-quality educational experiences.

The integration of multimodal generative AI technologies into storytelling systems significantly enhances both pedagogical effectiveness and learner engagement (Hsieh, 2021; Hu et al., 2022). Combining natural language processing, artificial intelligence, robotics, and educational theory allows humanoid robots to autonomously produce child-friendly narratives (Bravo et al., 2021). The resulting storytelling environment simultaneously increases language exposure during a formative period in children's development (Anderson et al., 2021; Chiang et al., 2020; Davies et al., 2021a, 2021b), thereby mitigating potential long-term educational and social repercussions linked to limited linguistic input (Hart & Risley, 1995).

Storytelling itself offers a means to captivate children's attention, immerse them in rich vocabulary and sentence structures, and convey essential cultural concepts (Kurniawan, 2021; Lee, 2022). Moreover, narratives broaden children's worldviews, helping them appreciate diverse customs, traditions, and values (Wright et al., 2007). A smart storytelling robot thus provides a continuous, interactive source of language input for children and supports parental efforts to sustain intellectual and emotional engagement, even in busy household contexts (Hu et al., 2022).

7.6.1 System Design and Architecture

From a design perspective, the robot's hardware and software work together to create an immersive environment. A dynamic display head, expressive facial features, and embedded sensors enable real-time responses to children's verbal and nonverbal cues, thereby establishing an emotionally resonant connection that bolsters attention and motivation (Hsieh, 2021). Through ChatGPT-based scripts, the robot aligns its speech, facial expressions, and gestures with the storyline and transitions seamlessly to tailored quizzes and games that cultivate comprehension, reasoning, and problem-solving skills (Davies et al., 2021a, 2021b). By aligning all of these elements with children's interests, such a storytelling robot offers a multifaceted learning platform that advances both language acquisition and cognitive growth (Anderson et al., 2021).

7.6.2 Hardware, Software, and Interactive Design

- Hardware**:** The robot employs the MQTT protocol to synchronize facial expressions, gestures, and multimedia content with the story's plot, creating a fully immersive experience.
- Software: A suite of generative AI tools (powered by API keys) enables multimodal input and output, allowing the robot to adjust its storytelling in real time.
- Interactive Design: The user-friendly chatbot interface, complemented by text-to-speech (TTS) and speech-to-text (STT) functionalities, promotes natural, intuitive communication. This design encourages children to actively participate in storytelling and learning, thereby enhancing both comprehension and cognitive development.

7.6.3 Key Features of the Smart Storytelling Robot

The smart storytelling robot exhibits several unique capabilities that set it apart from conventional storytelling devices:

1. AI-Driven Story Generation:
   Personalized stories are automatically created based on each child's interests and past interactions. By updating narratives over time, the system sustains long-term engagement.
2. Emotional Context Analysis:
   The robot uses AI algorithms to gauge story context, generating synchronized facial expressions, gestures, and multimedia elements that heighten immersion and emotional impact.
3. Interactive Question-And-Answer Sessions:
   Integrated into each story, Q&A segments and minigames deepen children's understanding, increasing engagement while fostering critical thinking and problem-solving.
4. Adaptive Learning Experience:
   By applying AI to adjust the difficulty and content of questions and games based on a child's performance, the robot tailors the learning journey to individual needs, thereby maximizing comprehension and retention.

The smart storytelling robot follows a systematic procedure that ensures a smooth, engaging user experience:

1. The multimodal intelligent storytelling robot introduces itself.
2. The robot asks the child what story they would like to hear.
3. The child states their preferred story.
4. The robot generates a story based on the child's request.
5. The robot, acting as a storyteller, narrates the story to the child.

6. The robot uses facial expressions to convey emotions that match the story's content.
7. The tablet displays text and images that correspond to the story's content.
8. The robot can provide multimedia content related to the story plot on the tablet, assisting the child in practicing storytelling in a contextualized environment.
9. The multimodal intelligent storytelling robot introduces itself.
10. The robot asks the child what story they would like to hear.
11. The child states their preferred story.
12. The robot generates a story based on the child's request.
13. The robot, acting as a storyteller, narrates the story to the child.
14. The robot uses facial expressions to convey emotions that match the story's content.
15. The tablet displays text and images that correspond to the story's content.
16. The robot can provide multimedia content related to the story plot on the tablet, assisting the child in practicing storytelling in a contextualized environment.

7.6.4 Expanded Potential and Multifunctional Applications

Although designed primarily for children's storytelling, the smart storytelling robot system offers broader applications across diverse demographic groups. Research on socially assistive robots suggests that AI-driven conversational features can alleviate loneliness and enhance mental well-being among older adults, particularly when human caregivers are limited or unavailable (Bemelmans et al., 2012; Leite et al., 2013). These findings align with Wada and Shibata's (2007) research, which highlights how emotionally responsive robots—capable of interpreting and adapting to user feedback—can sustain long-term engagement and provide practical support in daily life.

From a learning perspective, studies on AI-supported storytelling robots demonstrate significant benefits in motivation, language proficiency, and cognitive development (Hsieh, 2021; Lin et al., 2022). Extending these capabilities to older adults addresses both psychosocial needs (e.g., companionship) and cognitive stimulation. Personalized narratives, tailored to an individual's cultural background and life experiences, can aid in heritage preservation and identity reinforcement (Huston & Huston, 2024). Additionally, adaptive algorithms can be repurposed to modify story difficulty and question prompts, supporting rehabilitative and therapeutic interventions (Leite et al., 2013).

Ultimately, this AI-driven storytelling platform illustrates how a single technological framework can accommodate diverse user needs, from children developing language skills to seniors seeking cognitive and emotional engagement. This multifunctional approach aligns with the evolution of human–robot interaction (Alam, 2022), positioning the storytelling robot as both an educational tool for young learners and a supportive companion for older adults. By integrating AI-generated narratives, real-time adaptability, and emotionally responsive interactions, the robot

serves as a versatile model of interactive storytelling, enhancing learning experiences and quality of life across different life stages.

7.7 Conclusion

The culmination of this research presents a groundbreaking stride in the realm of foreign language education, made possible by the integration of advanced technologies in an Emulated Real-World Environment. The R&T learning system, at the heart of this innovation, harnesses the synergistic potential of Educational Robots and IoT-Enabled Tangible Objects to create a dynamic, immersive learning experience that closely mirrors real-world scenarios.

Central to the success of the R&T system is its ability to simulate realistic language use contexts, thereby addressing the prevalent challenges of immersion and contextual practice in language learning. The system's design, which facilitates interactions between learners and a humanoid robot, supplemented by tangible, IoT-enabled objects, has proven effective in enhancing key language skills. These include conversational proficiency, digital literacy, and a deeper understanding of physical concepts, all crucial components of comprehensive language education.

A noteworthy aspect of this research is the interdisciplinary collaboration that has driven the project. The convergence of expertise in content development, pedagogical strategies, and technological innovation has been instrumental in ensuring that the R&T system is not only educationally robust but also contextually relevant and technologically sophisticated. The positive outcomes observed in learners' enhanced language comprehension and communication skills are a testament to the effectiveness of this collaborative approach.

Furthermore, the extension of this system to early childhood language development through the use of AI-driven educational robots highlights the versatile application of this technology. It shows promise not only in traditional language-learning settings but also as a tool to bridge developmental language gaps in early learners.

In conclusion, this chapter represents a significant leap forward in the field of language education. By effectively merging robotics, IoT, and AI technologies, it has set a new benchmark for interactive, context-rich language learning. Looking ahead, the R&T learning system opens new avenues for innovation in educational technology, with the potential to revolutionize traditional language-learning methodologies and provide enriching, effective, and engaging language-learning experiences for learners of all ages.

Acknowledgements This research was supported by the National Science and Technology Council, Taiwan, under project numbers MOST 111-2410-H-003-028-MY3 and 109-2511-H-003-053-MY3. This work was financially supported by the "Institute for Research Excellence in Learning Sciences" of National Taiwan Normal University (NTNU) from The Featured Areas Research Center Program within the framework of the Higher Education Sprout Project by the Ministry of Education (MOE) in Taiwan."

References

Alam, A. (2022). Social robots in education for long-term human-robot interaction: Socially supportive behaviour of robotic tutor for creating robot-tangible learning environment in a guided discovery learning interaction. *ECS Transactions, 107*(1), 12389–12403. https://doi.org/10.1149/10701.12389ecst

Anderson, N. J., Graham, S. A., Prime, H., Jenkins, J. M., & Madigan, S. (2021). Linking quality and quantity of parental linguistic input to child language skills: A meta-analysis. *Child Development, 92*(2), 484–501. https://doi.org/10.1111/cdev.13508

Bemelmans, R., Gelderblom, G. J., Jonker, P., & de Witte, L. (2012). Socially assistive robots in elderly care: A systematic review into effects and effectiveness. *Journal of the American Medical Directors Association, 13*(2), 114–120. https://doi.org/10.1016/j.jamda.2010.10.002

Bravo, F. A., Hurtado, J. A., & González, E. (2021). Using robots with storytelling and drama activities in science education. *Education Sciences, 11*(7), 329. https://doi.org/10.3390/educsci11070329

Chen, J.-C. (2016). The crossroads of English language learners, task-based instruction, and 3D multi-user virtual learning in Second Life. *Computers & Education, 102*, 152–171. https://doi.org/10.1016/j.compedu.2016.08.004

Cheng, Y. W., Wang, Y., Cheng, Y. J., & Chen, N. S. (2022). The impact of learning support facilitated by a robot and IoT-based tangible objects on children's game-based language learning. *Computer Assisted Language Learning*, 1–32. https://doi.org/10.1080/09588221.2022.2152053

Chiang, Y.-H.V., Cheng, Y.-W., & Chen, N.-S. (2023). Improving language learning activity design through identifying learning difficulties in a platform using educational robots and IoT-based tangible objects. *Educational Technology & Society, 26*(3), 84–100. https://doi.org/10.30191/ETS.202307_26(3).0007

Chiang, Y. H. V., Zheng, Y. J., Cheng, Y. W., & Chen, N. S. (2020). Analyzing learners' English learning process involving educational robots and IoT based toys through the lens of zone of proximal development. In *IEEE 20th International Conference on Advanced Learning Technologies (ICALT)*, 213–215. https://doi.org/10.1109/ICALT49669.2020.00069

Davies, C., Hendry, A., Gibson, S. P., Gliga, T., McGillion, M., & Gonzalez-Gomez, N. (2021). Early childhood education and care (ECEC) during COVID-19 boosts growth in language and executive function. *Infant and Child Development*, *30*(4). https://doi.org/10.1002/icd.2241

Davies, C., Hendry, A., Gibson, S. P., Gliga, T., McGillion, M., & Gonzalez-Gomez, N. (2021). Early childhood education and care (ECEC) during COVID-19 boosts growth in language and executive function. *Infant and Child Development, 30*(4). https://doi.org/10.1002/icd.2241

Denden, M., Tlili, A., Chen, N.-S., Abed, M., Jemni, M., & Essalmi, F. (2024). The role of learners' characteristics in educational gamification systems: A systematic meta-review of the literature. *Interactive Learning Environments, 32*(3), 790–812. https://doi.org/10.1080/10494820.2022.2098777

Eseryel, D., Ifenthaler, D., & Ge, X. (2011). Alternative assessment strategies for complex problem solving in game-based learning environments. In D. Ifenthaler, Kinshuk, P. Isaias, D. G. Sampson & J. M. Spector (Eds.), *Multiple perspectives on problem solving and learning in the digital age* (pp. 159–178). Springer.

Fang, W.-C., Yeh, H.-C., Luo, B.-R., & Chen, N.-S. (2021). Effects of mobile-supported task-based language teaching on EFL students' linguistic achievement and conversational interaction. *ReCALL, 33*(1), 71–87. https://doi.org/10.1017/S0958344020000208

Fryer, L., & Carpenter, R. (2006). Emerging technologies bots as language learning tools. *Language Learning & Technology, 10*(3), 8–14. 10125/44068

Hart, B., & Risley, T. R. (1995). *Meaningful differences in the everyday experience of young American children*. Paul H. Brookes.

He, H., & Deng, Y. (2015). The mental lexicon and English vocabulary teaching. *English Language Teaching, 8*(7), 40–45. https://doi.org/10.5539/ELT.V8N7P40

Horwitz, E. K., Horwitz, M. B., & Cope, J. (1986). Foreign language classroom anxiety. *The Modern Language Journal, 70*(2), 125–132. https://doi.org/10.1111/j.1540-4781.1986.tb05256.x

Hsieh, J.-C. (2021). Digital storytelling outcomes and emotional experience among middle school EFL learners: Robot-Assisted versus powerpoint-assisted mode. *TESOL Quarterly, 55*(3), 994–1010. https://doi.org/10.1002/tesq.3043

Hsieh, W. M., Yeh, H. C., & Chen, N. S. (2023). Impact of a robot and tangible object (R&T) integrated learning system on elementary EFL learners' English pronunciation and willingness to communicate. *Computer Assisted Language Learning*, 1–26. https://doi.org/10.1080/09588221.2023.2228357

Hsu, F. H., Hsueh, Y. T., Chang, W. L., Lin, Y. T. R., Lan, Y. J., & Chen, N. S. (2022). An innovative multimodal learning system based on robot and tangible objects for Chinese numeral-classifier-noun phrase learning. In *2022 International Conference on Advanced Learning Technologies (ICALT).* IEEE. https://doi.org/10.1109/icalt55010.2022.00072

Hu, C. C., Chen, M. H., Yuadi, I., & Chen, N. S. (2022, February). The effects of constructing robot-based storytelling system on college students' computational thinking skill and technology comprehension. In *24th International Conference on Advanced Communication Technology (ICACT),* 496–500. https://doi.org/10.23919/ICACT53585.2022.9728803

Hu, C. C., Tseng, H. T., Chen, M. H., Imm, A. G. P., & Chen, N. S. (2020). Comparing the effects of robots and IoT objects on STEM learning outcomes and computational thinking skills between programming-experienced learners and programming-novice learners. In *IEEE 20th International Conference on Advanced Learning Technologies (ICALT)*, 87–89. https://doi.org/10.1109/ICALT49669.2020.00033

Hu, C. C., Yang, Y. F., & Chen, N. S. (2023). Human–Robot interface design–The 'Robot with a Tablet' or 'Robot only', which one is better? *Behaviour & Information Technology, 42*(10), 1590–1603. https://doi.org/10.1080/0144929X.2022.2093271

Hutson, J., & Hutson, P. (2024). Storytelling. In *Inclusive smart museums: Engaging Neurodiverse audiences and enhancing cultural heritage* (pp. 49–84). Springer Nature Switzerland.

Hwang, G. J., & Chen, N. S. (2023). Editorial position Paper: Exploring the potential of generative artificial intelligence in education: applications, challenges, and future research directions. *Educational Technology & Society, 26*(2). https://doi.org/10.30191/ETS.202304_26(2).0014

Ifenthaler, D., Kinshuk, I., & P., Sampson, D. G., & Spector, J. M. (Eds.). (2011). *Multiple perspectives on problem solving and learning in the digital age.* Springer.

Konijn, E. A., Jansen, B., Mondaca Bustos, V., Hobbelink, V. L., & Preciado Vanegas, D. (2022). Social robots for (second) language learning in (migrant) primary school children. *International Journal of Social Robotics*, 1–17. https://doi.org/10.1007/s12369-021-00824-3

Kory Westlund, J. M., Jeong, S., Park, H. W., Ronfard, S., Adhikari, A., Harris, P. L.,... & Breazeal, C. L. (2017). Flat vs. expressive storytelling: Young children's learning and retention of a social robot's narrative. *Frontiers in Human Neuroscience*, *11*, 295. https://doi.org/10.3389/fnhum.2017.00295

Kurniawan, M. (2021). Digital storytelling: Teachers' guide to attract children's interest and motivation in kindergarten's English language learning. *Satya Widya, 37*(1), 16–24. https://doi.org/10.24246/j.sw.2021.v37.i1.p16-24

Lee, J. (2022). A meta-analysis of the effects of storytelling on English language learning. *Modern English Education, 23*(2), 33–51. https://doi.org/10.18095/meeso.2022.23.2.33

Leite, I., Martinho, C., & Paiva, A. (2013). Social robots for long-term interaction: A survey. *International Journal of Social Robotics, 5*(2), 291–308. https://doi.org/10.1007/s12369-013-0178-y

Lin, V., Yeh, H.-C., & Chen, N.-S. (2022). A systematic review on oral interactions in robot-assisted language learning. *Electronics, 11*(2), 290. https://doi.org/10.3390/electronics11020290

Lin, V., Yeh, H.-C., Huang, H.-H., & Chen, N.-S. (2021). Enhancing EFL vocabulary learning with multimodal cues supported by an educational robot and an IoT-Based 3D book. *System, 104*, 102691–102691. https://doi.org/10.1016/j.system.2021.102691

Mayer, R. E. (2005). Cognitive theory of multimedia learning. *The Cambridge Handbook of Multimedia Learning*, *41*(1), 31–48. https://doi.org/10.1017/CBO9780511816819.004

Milton, J., Jonsen, S., Hirst, S., & Lindenburn, S. (2012). Foreign language vocabulary development through activities in an online 3D environment. *The Language Learning Journal, 40*(1), 99–112. https://doi.org/10.1080/09571736.2012.658229

Peura, L., Mutta, M., & Johansson, M. (2023). Playing with pronunciation: A study on robot-assisted French pronunciation in a learning game. *Nordic Journal of Digital Literacy, 2*, 100–115. https://doi.org/10.18261/njdl.18.2.3

Renandya, W. A. (2013). Essential factors affecting EFL learning outcomes. *English Teaching, 68*(4), 23–41. https://doi.org/10.15858/engtea.68.4.201312.23

Sweller, J. (2011). Cognitive load theory. In *Psychology of learning and motivation* (Vol. 55, pp. 37–76). Academic Press.

Tseng, W.-T., Liu, Y.-T., Hsu, Y.-T., & Chu, H.-C. (2021). Revisiting the effectiveness of study abroad language programs: A multi-level meta-analysis. *Language Teaching Research, 28*(1), 1–45. https://doi.org/10.1177/1362168820988423

Tutal, Ö. (2023). Does context-based learning increase academic achievement and earning retention?: A review based on meta-analysis. *Journal of Practical Studies in Education, 4*(5), 1–16.

Wada, K., & Shibata, T. (2007). Living with seal robots—Its sociopsychological and physiological influences on the elderly at a care house. *IEEE Transactions on Robotics, 23*(5), 972–980. https://doi.org/10.1109/TRO.2007.906261

Wahl, H., & Winiwarter, W. (2012). A context-related vocabulary trainer in the integrated intelligent computer-assisted language learning (iiCALL) environment. In *Proceedings of the 14th International Conference on Information Integration and Web-Based Applications & Services.* ACM. https://doi.org/10.1145/2428736.2428802

Wenas, I., Liando, N. V. F., & Rorimpandey, R. (2023). Task based learning as an active retrieval approach in the perception of EFL students: A case study at bridge education Manado. *Journal of English Culture, Language, Literature and Education, 11*(2), 189–209.

Wright, C., Bacigalupa, C., Black, T., & Burton, M. (2007). Windows into children's thinking: A guide to storytelling and dramatization. *Early Childhood Education Journal, 35*(4), 363–369. https://doi.org/10.1007/s10643-007-0189-0

Yang, Y. F., Goh, A. P. I., Hu, C. C., Chen, N. S., & Lai, S. C. (2024). Digital game-playing to enhance English vocabulary and content learning among anxious foreign language students. *Computer Assisted Language Learning*, 1–28. https://doi.org/10.1080/09588221.2024.2371383

Zhao, Z. (2021). Virtual reality language lab of second language learning based on J Bruner's discovery learning and self-efficacy theory—Literature review and solution improvement. *Journal of Innovation and Social Science Research ISSN, 2591*, 6890. https://doi.org/10.53469/jissr.2021.08(07).30

Dr. Nian-Shing Chen is a Chair Professor in the Institute for Research Excellence in Learning Sciences and Program of Learning Sciences at National Taiwan Normal University, Taiwan. His research areas include robot-assisted language learning, embodied cognition, Internet of Things (IoT), Blockchain, and the application of educational robots in education. Professor Chen received the National Outstanding Research award three times in 2008, 2011, and 2015. He is ranked in second place on the Top Authors list by article count in Computers and Education. He is ranked number 4 on the top scholar list in the educational technology field based on an article published in the British Journal of Educational Technology. One of his papers published in Innovations in Education and Teaching International was awarded as the top-cited article in 2010. Professor Chen is ranked third in the list of Social Sciences and Humanities Scientists in Taiwan. Based on AD Scientific Index 2023, Prof. Chen is ranked number 8 in Asia and is ranked number 103 in the world in Education Research. He is also listed as one of Scopus World Top 2% Scientists (2021)

(Ioannidis, 2022). Professor Chen previously served as the Editor-in-Chief for Educational Technology and Society (ET&S) journal. He is a golden core member of IEEE, a lifelong member of ACM, and the former Chair of the IEEE Technical Committee ons Learning Technology.

Professor Hui-Chin Yeh Distinguished Professor and Dean of International Affairs, is acknowledged as one of the World's Top 2% Scientists. Professor Yeh's scholarly pursuits encompass a broad spectrum of cutting-edge educational technologies and methodologies. Her work significantly contributes to the enhancement of language teaching and learning through the adoption of advanced tools such as robotics, augmented reality (AR), and virtual reality (VR). She also explores the incorporation of multimodality in supporting content creation, enriching the learning experience beyond traditional text-based approaches. Additionally, she pioneers in facilitating telecollaboration across international borders, fostering language exchange and cultural understanding through virtual partnerships and exchanges. These diverse research interests underscore her commitment to expanding the scope and efficacy of multiliteracies pedagogy. Professor Yeh's academic journey is distinguished by the publication of over 70 research papers, with a remarkable 85% of these contributions being featured in SSCI-indexed journals, reflecting the significant impact and high quality of her research. Her scholarly reputation is further enhanced by her being honored as a recipient of the Distinguished Young Scholar Awards from the National Science Technology and Council on four occasions, in the years 2017, 2011, 2020, and 2023. These awards are emblematic of the continuous recognition of her exemplary research and her ongoing commitment to academic excellence.

Chapter 8
AI for Language Learning: Opportunities and Challenges

Jo-Chi Hsiao, Shiowyun Chang, and Jason S. Chang

Abstract The advent of the latest artificial intelligence (AI) technology, such as large language models (LLM), and AI services, such as ChatGPT, has the potential to reform the way we learn a language and the way we study language learning. The widespread interest in ChatGPT shows text writing, rephrasing, correction, and even translation have become easily accessible to people from all walks of life. With its capacity to simulate complex human-like interactions and offer immediate, nuanced feedback, AI serves as a potent tool for enhancing the effectiveness of language acquisition. Given these opportunities and demonstrated efficacy, AI-mediated language learning (AIM-LL) will be an active research area for a long time to come. This chapter provides an overview of the technology behind ChatGPT to enable a more comprehensive engagement with this AI system. We also include case studies on language learning using ChatGPT and other AI tools to showcase opportunities in AI-mediated language learning and research. Yet, as promising as these prospects are, they do not come without challenges. These include ethical concerns, the digital divide, and the potential loss of the human touch in language education. This chapter specifies some of these challenges, offering a balanced perspective on the role of AI in language learning. Finally, we point out future technological innovations following the success of ChatGPT to provide insight into the trajectory of AI development. This topic is relevant to language learners, educators, researchers, technologists, and policymakers interested in the future of language education.

Keywords Computer-assisted language learnings · Intelligent tutoring systems · Artificial intelligence · Generative AI · English teaching

J.-C. Hsiao
National Hsinchu Senior High School, Hsinchu, Taiwan
e-mail: annahsiao@nlplab.cc

S. Chang
Hsinchu Chien Kung Senior High School, Hsinchu, Taiwan
e-mail: julyun@cksh.hc.edu.tw

J. S. Chang (✉)
National Tsing Hua University, Hsinchu, Taiwan
e-mail: jason@nlplab.cc

Y. Lan et al. (eds.), *AI-Mediated Language Education in the Metaverse Era*, Chinese Language Learning Sciences, https://doi.org/10.1007/978-981-95-0245-5_8

8.1 Introduction

The year 2023 ushered in an unprecedented era of technological innovation, with artificial intelligence (AI) emerging as a transformative force across many diverse domains. For example, in the realm of TESOL (Teaching English to Speakers of Other Languages), the integration of AI-powered tools has sparked a paradigm shift, and redefined conventional teaching and learning. This chapter aims to prompt educators and scholars to closely scrutinize the potential benefits, drawbacks, and nuanced dynamics that emerging AI (notably, ChatGPT) brings into the TESOL landscape. With that in mind, we include three main sections collectively examining the impact of AI on English language teaching and learning in the context of Taiwan, with a focus on using ChatGPT. The exploration ranges from crafting golden prompts () for writing tools, to experimenting with ChatGPT to enrich students' educational experiences, and finally to a mixed-method study for comprehensively examining the positive and negative effects of using ChatGPT from both teachers' and students' perspectives.

The three sections contribute to this chapter in different ways. In the section *Golden prompts for writing tools*, we present our insights of developing writing assistance tools based on ChatGPT, underscoring its potential to revolutionize the learning experience of English writing. In the section *AI-enhanced English teaching and learning*, we detail the roles of AI tools in reducing teachers' workload and enhancing efficiency with many examples of student works. In the section *ChatGPT, teachers, and their students*, we describe a mixed-methods study, unveiling the advantages of AI integration in English classrooms and delineating pertinent disadvantages for both teachers and learners.

8.2 Golden Prompts for Writing Tools

Leading up to the year 2023, the world was captivated by a series of newly released AI tools which have the potential to reshape the landscape of language learning. Notably, ChatGPT engages users in multi-turn dialogues, answering questions and responding to requests for information and assistance. Unlike earlier AI technologies, these responses are more effective, particularly when "prompts" (i.e., the input provided to an AI system, in the form of questions, requests, or instructions) are well thought out, as highlighted by recent research in Prompt Engineering (e.g., Meskó, 2023; White et al., 2023).

This section presents insights from our experience in crafting so-called "golden prompts" (i.e., …PLEASE DEFINE) to leverage ChatGPT for developing writing tools, which offer suggestions to enhance grammaticality, logic, word choice, collocation, and organization, as well as comprehensive explanations.

In a 2023 North Bay Python Workshop, Simon Willison said: "*About three years ago, aliens landed on Earth, and they handed over a USB hard stick, and they*

disappeared. And since then, we've been poking the thing they gave us with a stick trying to figure out what it does, how it works, and what's more it can do" (www.youtube.com/watch?v=h8Jth_ijZyY). In a nutshell, this thing is a file containing numbers, or a program (function) that comprises an exceedingly large language model (LLM). In its simple, historical form, a language model (LM) predicts the next word after two preceding words (context), after training (i.e., counting) tens of millions of words of textual data. However, when LMs scaled up and became LLM, the context became larger and larger as the training data ran into terabytes. Suddenly surprising and unexpected abilities emerged. Kevin Kelley famously said: "*Any sufficiently advanced technology is indistinguishable from magic*". We could say ChatGPT is a magical gift from aliens, because we are still continuously surprised by more and more things we can do using LLMs, ChatGPT in particular.

For those who are not familiar with the history of Natural Language Processing (NLP) and Artificial Intelligence (AI) in general, a brief summary of historic events leading to this magic moment is in order.

In 1989, Tim Berners-Lee, a British scientist, put forward the proposal of World Wide Web (WWW) while working at CERN (Conseil Européen pour la Recherche Nucléaire, or European Council for Nuclear Research in English). WWW (or the Internet) is the reason that we can gather and store a mind-boggling amount of data to train these LLMs. Evidently, more data is better data as Church and Mercer (1993) so aptly put it. Apparently, there are no other data sources that have more data than the Internet.

However, quality does matter to a large extent, as prompt engineering studies revealed. The hallmark of a high-quality, authoritative information source is the Encyclopedia Britannica (https://www.britannica.com/). In 2001, Jimmy Wales and Larry Sanger founded Wikipedia as a new, crowd-sourced encyclopedia. Researchers soon found that Wikipedia surpassed the Encyclopedia Britannica in quantity, and even quality. Wikipedia was at first the core, reliable source of LLM, despite its small size as compared to the crawled datasets from the Internet, such as the News Crawl for machine translation research (www.statmt.org/wmt13/translation-task.html). A crawled dataset (or Crawl) is created by designing and running a software program (i.e., "crawler") to go from Web pages to Web pages via hyperlinks and download the textual and other types of information on the pages.

In 2007, Gil Elbaz founded Common Crawl Project which became the main body of LLM training materials. The purpose of Common Crawl is to counter the monopoly of search engine companies having access to the data on the Internet.

Meanwhile, the other branches of Artificial Intelligence, notably Image Processing (IP), have been very successful, mainly because of ImageNet (Deng et al., 2009) and the image representation as a matrix for efficient parallel processing using Graphics Processing Units (GPU). With the available large-scale textual datasets (much bigger than ImageNet), NLP now needs a good vector-based representation of data (words, sentences, and text) like IP to garner the same successes that IP has already achieved.

This moment happened six years later, as Mikolov, Chen, Corrado, Dean (2013) at Google published the word2vec paper describing a two-layer neural network (NN) for turning words in a large newspaper dataset (Google News) into vectors.

The ensuing years saw more and more innovative NNs such as RNN (Recursive Neural Network, which can accept a flexible number of input words), and LSTM (Long Short-Term Memory, a type of enhanced RNN designed to capture long-term dependencies in a sentence for better results), with an increasing number of layers and complicated topology. This trend led Vaswani et al. at Google Brain and Google Research to develop the best NN architecture ever, the Transformer in 2017. Up to this point, all NNs have been used in a setting of Supervised Machine Learning (ML), where a model is "trained" with a set of corresponding input and output pairs for a specific task, such as Question Answering or Machine Translation.

Most of the Web-borne data are raw with no annotation for any task-specific information (supervision). In the spirit of unsupervised or self-supervised learning, in 2018, Devlin, Chang, Lee, and Toutanova at Google AI developed BERT (Bidirectional Encoding Transformer). Within one year, more than 10,000 papers were published citing BERT in their works that successfully used BERT as a pre-training model to fine-tune models for specific tasks with limited annotated training data. The success of BERT lies in using raw sentences to generate automatically a large amount of training data for two tasks: fill-in-the-blank, and next-sentence-prediction. In the next year, 2019, Colin Raffel et al. at Google Brain introduced the T5 model taking the BERT approach to another level.

Interestingly a year earlier, researchers at OpenAI started to develop GPT-1, as the first iteration in the GPT series. GPT-1 was revolutionary for its time, introducing the transformer architecture for effective language modeling and generation. With a stronger foundation model for text generation (Generative AI), the GPT series later incorporated mechanisms to handle conversation, leading to ChatGPT in late 2022.

ChatGPT is specifically optimized for generating coherent and contextually relevant text in a conversational style, making it well-suited for applications like chatbots or interactive text-based AI systems. T5, on the other hand, is a more general-purpose model that treats all text-based tasks as a problem of converting one form of text into another, making it highly versatile for a wide range of natural language processing tasks.

The technological breakthrough of ChatGPT brings us great opportunities to develop new writing tools, with added assistive functions, such as more effective grammatical error correction, word choice improvement with explanation, simple rewriting, and improving the given text. Additionally, we could also use ChatGPT to generate example sentences directly used by teachers and students, or alternatively in training an AI model.

These opportunities do incur some risks. First, using ChatGPT naively to improve student essays with very general prompts runs the risk of obtaining over-corrected results. If we ask for an explanation of the correction, the explanation could be too superficial to be the useful. And, often the evidence of do's and don'ts tend to be qualitative rather than quantitative. Sentences, paragraphs, and text can be improved in a way not intended if we prompt ChatGPT without giving specific directions toward improvement.

In our quest to develop writing assistance tools using ChatGPT for English learners, we first experimented with a traditional Grammar Checker, Linggle Write,

and simple ChatGPT prompts to see how they identify grammatical errors and provide suggestions (see Fig. 8.1). Then, we investigated how to use ChatGPT to explain the errors and suggestions. Subsequently, we explored using more elaborate prompts to show how Prompt Engineering can be applied in language learning. The prompts we have experimented with can be used by individual learners for self-editing; alternatively, these specific prompts can be built into a grammatical error correction system for improving performance and expanding functionality.

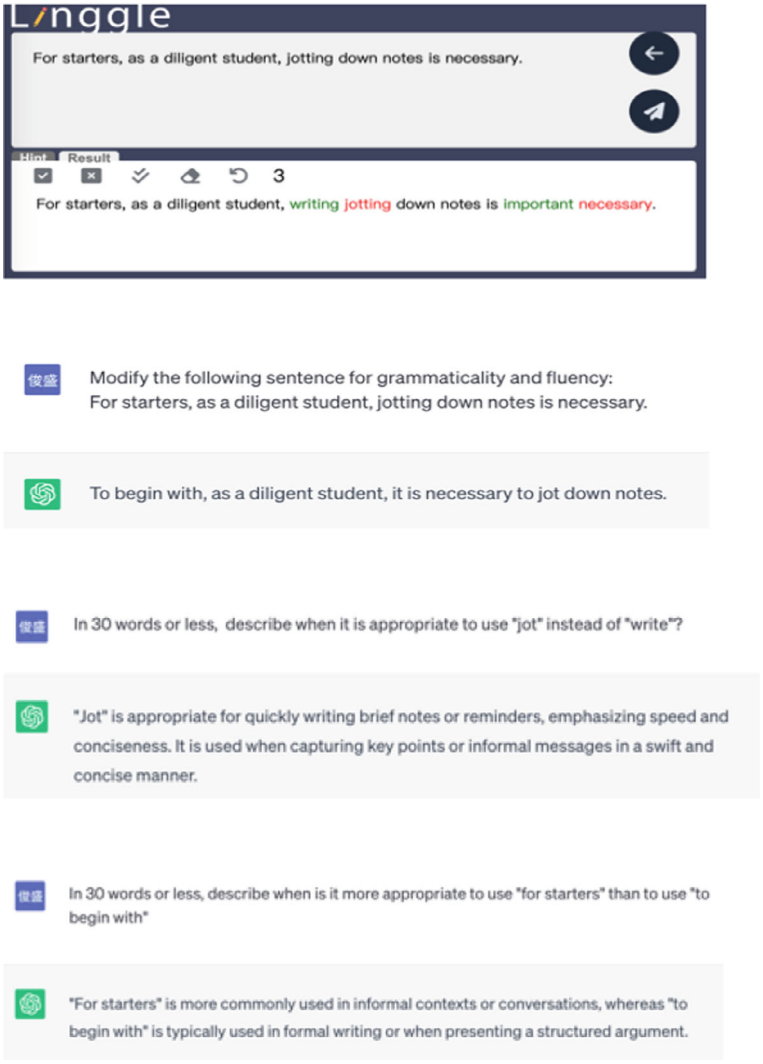


Fig. 8.1 Identifying grammatical errors and providing suggestions with Linggle Write and ChatGPT

(1) Consider, for example, a sentence written by a high school student in Taiwan and the editorial results obtained using Linggle Write and ChatGPT: For starters, as a diligent student, jotting down notes is important. (Original)
(2) For starters, as a diligent student, writing down notes is necessary. (Linggle Write)
(3) To begin with, as a diligent student, it is necessary to jot down notes. (ChatGPT)
(4) "Jot" is used to emphasize quick notes and reminders, emphasizing speediness and conciseness. (ChatGPT)
(5) "For starters" is more often used in informal contexts and conversations, whereas "To begin with" is typically used in formal writing or presenting a structural argument. (ChatGPT)

Sentence (2) returned by Linggle Write shows that a state-of-the-art grammatical error correction system (GEC) can identify syntax errors as well as word choice errors quite successfully, whereas generative AI tools such as ChatGPT are more capable of returning phrasal and structural edits. But the potential of Gen AI tools does not stop here with a simple prompt like "Modify the following sentence for grammaticality and fluency". Most strikingly, ChatGPT can be prompted to explain edits, of its own or by other GEC systems (as in Examples (4) and (5)).

We can opt for receiving small, word-level, conservative edits from ChatGPT by using the verb "proofread", instead of "modify" or "improve". For example, Prompt (6) led ChatGPT to respond with word-level edits like "We must acquire knowledge every day".

(6) Proofread the following text: "We must learn knowledge every day".

Additionally, we have experimented with the following prompts based on research findings on Prompt Engineering:

(7) Act as if you are a seasoned English teacher to help improve my text by making the sentence more grammatical and logical, with better word choices and idiomatic collocations, in Standard English use. Now, proofread the sentence: "It is difficult to reach abandoned places such as small country villages. Her car was involved in big accident". Mark each identified error to be removed or replaced with square brackets (i.e., "[[", "]]"), and then insert the suggested word to replace the error or to add in with curly brackets (i.e., "{{", "}}") right after each error.

With this chain-of-thoughts prompt (i.e., step by step), we are able to obtain the following from ChatGPT:

Certainly, let's address the sentence:

> "It is difficult to reach abandoned places such as small country villages. Her car was involved in big accident".
>
> 1. The term "abandoned places" might not be the best descriptor for "small country villages". Not all small country villages are abandoned.
> 2. The phrase "big accident" is missing an article, and "big" can be replaced with a more appropriate adjective.

Here's the proofread sentence:

"It is difficult to reach [[abandoned places]]{{remote locations}} such as small country villages. Her car was involved in [[big]]{{a serious}} accident".

In conclusion, ChatGPT is indeed an alien technology, and its nature, limitations, and scope have yet to be fully explored. ChatGPT can be used to develop teaching tools for vocabulary, grammar, discourse, and pragmatics. The essential point is that creating well-designed built-in prompts to embed in writing tools based on research leads to better writing tools.

8.3 AI-Enhanced English Teaching and Learning

According to data from McKinsey & Company (Bryant, 2020), teachers in Canada, Singapore, the United Kingdom, and the United States work an average of 50 h per week. As illustrated in Fig. 8.2, teachers allocate approximately 10.5 h per week to preparation, 6.5 h to evaluation and feedback, and 3 h to professional development. Administrative tasks account for an additional 5 h. Collectively, these activities total 25 h per week. As Bill Gates has observed, "The age of AI has begun". This raises the question: Can AI tools enable teachers to reduce the time spent on these tasks while maintaining or even enhancing the quality of teaching? The following section discussing one of a teacher-researcher's experiences with AI-powered tools in the classroom may shed some light on the question.

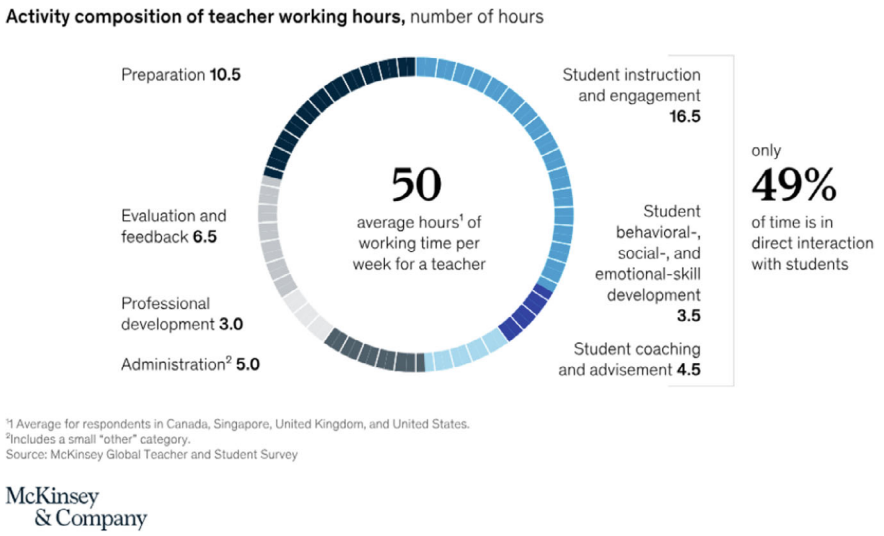


Fig. 8.2 Teachers' average weekly work hours and activities in Canada, Singapore, the United Kingdom, and the United States, Bryant 2020

This section is divided into two parts. The first part examines the integration of ChatGPT and other artificial intelligence (AI) tools within educational settings. The second part details an AI-driven semester project undertaken by students in one of the author's elective courses.

8.3.1 ChatGPT and Other AI Tools in Education

The teacher-researcher initially used ChatGPT, an AI tool that Bill Gates has claimed "will change our world", to improve her students' writing skills, identify errors, and provide practice exercises. The instructional process began with the presentation of a tutorial slide deck based on a free online course by Haishuo Lee, the superintendent of Wagor International School. Students were guided through a series of steps, including creating a ChatGPT account, logging in, and utilizing the tool to evaluate their essays, identify errors, and then recorded the feedback from GPT on their worksheets. Subsequently, they used GPT to generate practice questions targeting their specific errors and received revised versions of their essays from the AI (Fig. 8.3). The final component of the tutorial emphasized ethical considerations in the use of ChatGPT, ensuring that students understood the importance of using this AI tool responsibly (Fig. 8.4). Given that AI tools like ChatGPT can be both advantageous and potentially problematic, it is crucial for educators to actively monitor and guide students' use of these tools and address any emerging issues related to dependency.

For most of the students, this was their first experience with ChatGPT. When submitting prompts, reviewing the responses, and completing worksheets that highlighted the revisions made by the AI, they were visibly impressed by the speed of ChatGPT's responses. The classroom was filled with expressions of astonishment, and student feedback was overwhelmingly positive. For example, one of the students wrote as feedback in the worksheet (Fig. 8.6): "Although some people use ChatGPT for assignments or other improper purposes, I find it a great tool for improving basic English skills. It quickly identifies the errors I've made in grammar and vocabulary and provides valuable revisions. While it can't replace real conversations, ChatGPT is my best option when a teacher isn't available".

This experience significantly broadened the teacher's pedagogical approach, leading to the integration of ChatGPT in various educational contexts. For instance, ChatGPT was used to generate a slide deck on identifying noun clauses and adjective clauses, as well as a list of prompts designed to support students in enhancing their English skills (Figs. 8.5 and 8.6). Additionally, ChatGPT was utilized to generate a worksheet for creative writing (Fig. 8.7) and will be discussed in greater detail below. Beyond worksheets, ChatGPT was instrumental in developing various test items, including those for vocabulary, cloze exercises, and reading comprehension. While the results were not infallible and occasionally contained inaccuracies, thus necessitating further review and revision, ChatGPT nonetheless proved to be a valuable tool in reducing the time and effort required for preparing teaching materials.

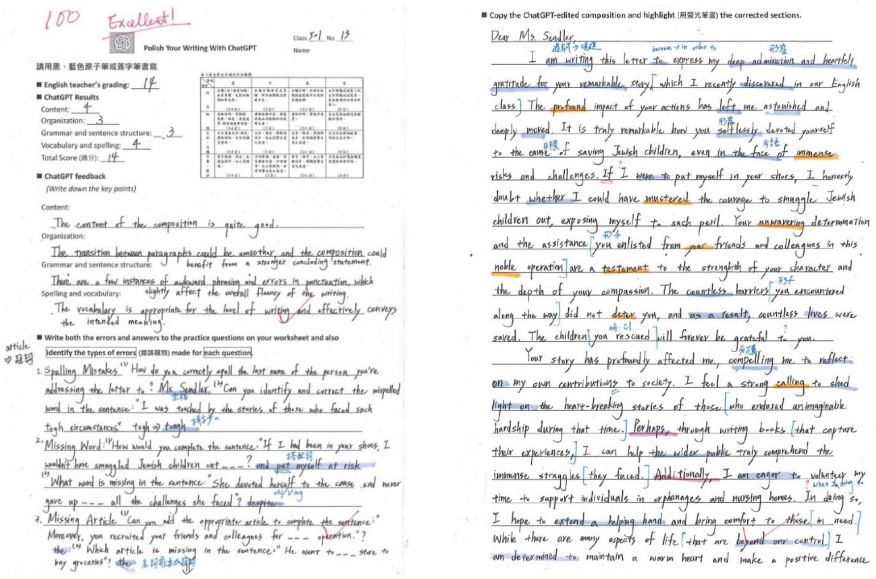


Fig. 8.3 A student's completed worksheet recording ChatGPT's feedback on her writings. In the worksheet, she recorded ChatGPT's feedback on content, organization, grammar, sentence structure, spelling, and vocabulary, all of which align with GSAT standards. She then identified the types of errors she'd made and wrote her answers to the practice questions provided by ChatGPT. Finally, she copied the ChatGPT-edited version of her writing and highlighted the correct sections

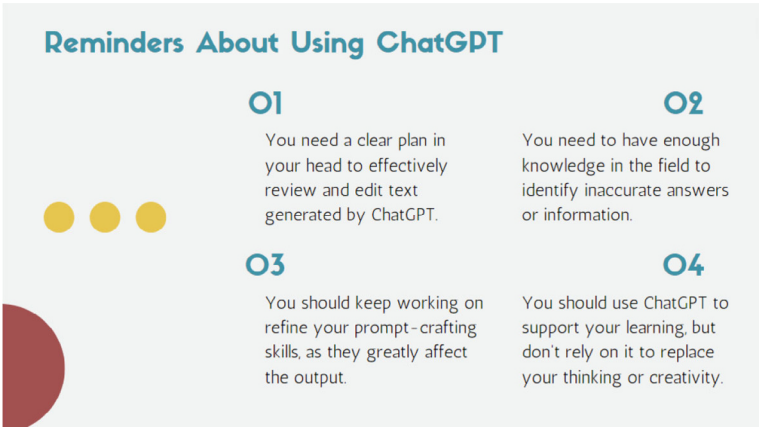
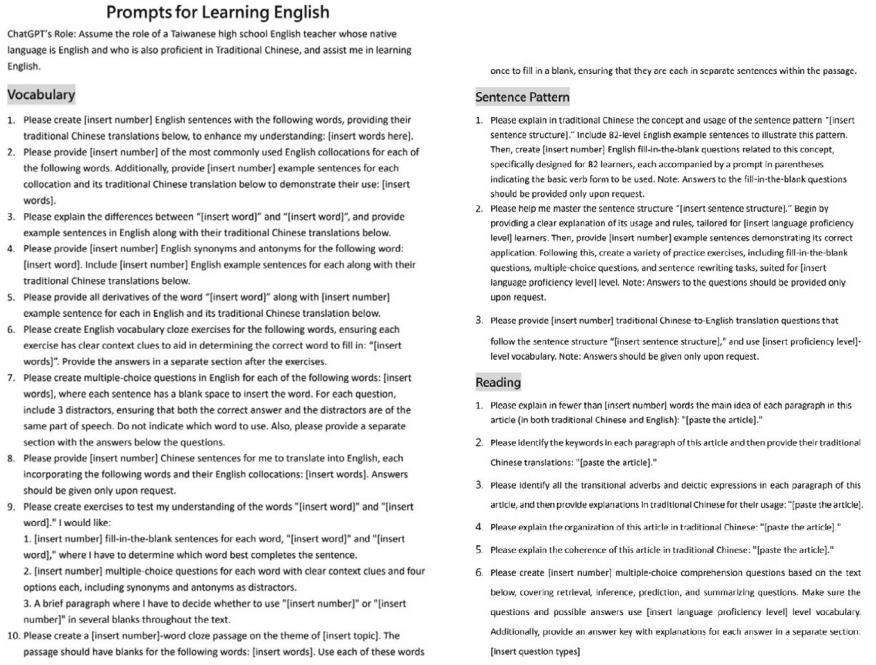


Fig. 8.4 A slide designed to reinforce proper usage of ChatGPT for students



Prompts for Learning English

ChatGPT's Role: Assume the role of a Taiwanese high school English teacher whose native language is English and who is also proficient in Traditional Chinese, and assist me in learning English.

Vocabulary

1. Please create [insert number] English sentences with the following words, providing their traditional Chinese translations below, to enhance my understanding: [insert words here].
2. Please provide [insert number] of the most commonly used English collocations for each of the following words. Additionally, provide [insert number] example sentences for each collocation and its traditional Chinese translation below to demonstrate their use: [insert words].
3. Please explain the differences between "[insert word]" and "[insert word]", and provide example sentences in English along with their traditional Chinese translations below.
4. Please provide [insert number] English synonyms and antonyms for the following word: [insert word]. Include [insert number] English example sentences for each along with their traditional Chinese translations below.
5. Please provide all derivatives of the word "[insert word]" along with [insert number] example sentence for each in English and its traditional Chinese translation below.
6. Please create English vocabulary cloze exercises for the following words, ensuring each exercise has clear context clues to aid in determining the correct word to fill in: "[insert words]". Provide the answers in a separate section after the exercises.
7. Please create multiple-choice questions in English for each of the following words: [insert words], where each sentence has a blank space to insert the word. For each question, include 3 distractors, ensuring that both the correct answer and the distractors are of the same part of speech. Do not indicate which word to use. Also, please provide a separate section with the answers below the questions.
8. Please provide [insert number] Chinese sentences for me to translate into English, each incorporating the following words and their English collocations: [insert words]. Answers should be given only upon request.
9. Please create exercises to test my understanding of the words "[insert word]" and "[insert word]." I would like:
   1. [insert number] fill-in-the-blank sentences for each word, "[insert word]" and "[insert word]," where I have to determine which word best completes the sentence.
   2. [insert number] multiple-choice questions for each word with clear context clues and four options each, including synonyms and antonyms as distractors.
   3. A brief paragraph where I have to decide whether to use "[insert number]" or "[insert number]" in several blanks throughout the text.
10. Please create a [insert number]-word cloze passage on the theme of [insert topic]. The passage should have blanks for the following words: [insert words]. Use each of these words once to fill in a blank, ensuring that they are each in separate sentences within the passage.

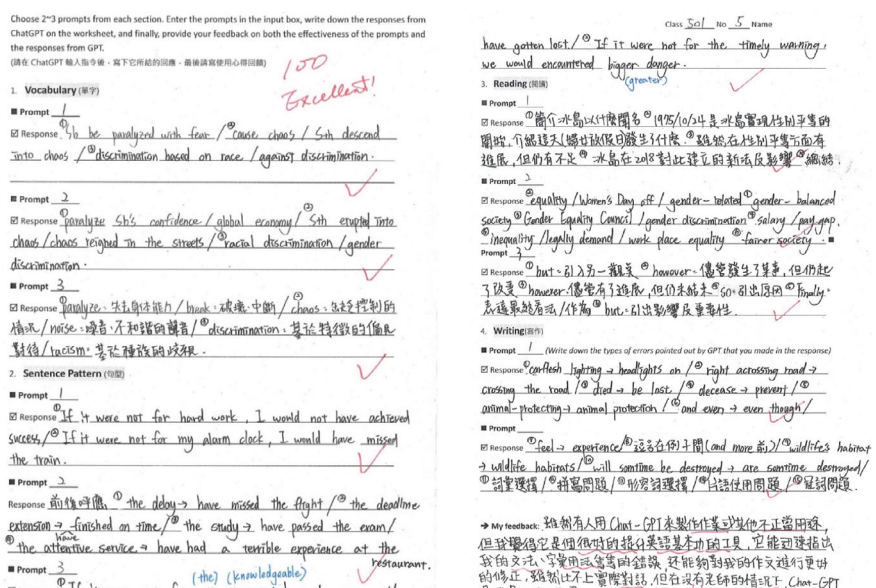
Sentence Pattern

1. Please explain in traditional Chinese the concept and usage of the sentence pattern "[insert sentence structure]." Include B2-level English example sentences to illustrate this pattern. Then, create [insert number] English fill-in-the-blank questions related to this concept, specifically designed for B2 learners, each accompanied by a prompt in parentheses indicating the basic verb form to be used. Note: Answers to the fill-in-the-blank questions should be provided only upon request.
2. Please help me master the sentence structure "[insert sentence structure]." Begin by providing a clear explanation of its usage and rules, tailored for [insert language proficiency level] learners. Then, provide [insert number] example sentences demonstrating its correct application. Following this, create a variety of practice exercises, including fill-in-the-blank questions, multiple-choice questions, and sentence rewriting tasks, suited for [insert language proficiency level] level. Note: Answers to the questions should be provided only upon request.
3. Please provide [insert number] traditional Chinese-to-English translation questions that follow the sentence structure "[insert sentence structure]," and use [insert proficiency level]-level vocabulary. Note: Answers should be given only upon request.

Reading

1. Please explain in fewer than [insert number] words the main idea of each paragraph in this article (in both traditional Chinese and English): "[paste the article]."
2. Please identify the keywords in each paragraph of this article and then provide their traditional Chinese translations: "[paste the article]."
3. Please identify all the transitional adverbs and deictic expressions in each paragraph of this article, and then provide explanations in traditional Chinese for their usage: "[paste the article]."
4. Please explain the organization of this article in traditional Chinese: "[paste the article]."
5. Please explain the coherence of this article in traditional Chinese: "[paste the article]."
6. Please create [insert number] multiple-choice comprehension questions based on the text below, covering retrieval, inference, prediction, and summarizing questions. Make sure the questions and possible answers use [insert language proficiency level] level vocabulary. Additionally, provide an answer key with explanations for each answer in a separate section: [insert question types]

Fig. 8.5 Prompts designed to assist students in enhancing their English skills with the aid of ChatGPT, including the categories of vocabulary, sentence patterns, and reading



Choose 2~3 prompts from each section. Enter the prompts in the input box, write down the responses from ChatGPT on the worksheet, and finally, provide your feedback on both the effectiveness of the prompts and the responses from GPT.
(請在 ChatGPT 輸入指令後，寫下它所給的回應，最後請寫使用心得回饋)

100 Excellent!

1. **Vocabulary** (單字)

■ Prompt 1
☑ Response ①to be paralyzed with fear / ②cause chaos / 5th descend into chaos / ③discrimination based on race / against discrimination.

■ Prompt 2
☑ Response ①paralyze 5th's confidence / global economy / 5th erupted into chaos / chaos reigned in the streets / ②racial discrimination / gender discrimination.

■ Prompt 3
☑ Response ①paralyze: 失去身体能力 / freak: 破壞、中斷 / ②chaos: 造成控制的情況 / noise: 噪音、不和諧的聲音 / ③discrimination: 基於特徵的偏見對待 / racism: 基於種族的歧視.

2. **Sentence Pattern** (句型)

■ Prompt 1
☑ Response ①If it were not for hard work, I would not have achieved success. / ②If it were not for my alarm clock, I would have missed the train.

■ Prompt 2
Response 前往時應 ①the delay → have missed the flight / ②the deadline extension → finished on time / ③the study → have passed the exam / ④the attentive service → have had a terrible experience at the restaurant.

■ Prompt 3
☑ Response ①If it were not for an acquainted guide, we would have gotten lost. / ②If it were not for the timely warning, we would encountered bigger danger. (greater)

Class Sa1 No. 5 Name

3. **Reading** (閱讀)

■ Prompt 1
☑ Response ①簡介冰島以什麼開名 ②1975/10/24 是冰島當現性別平等的開始，介紹這天婦女放假日發生了什麼 ③雖然在性別平等方面有進展，但仍有不足 ④冰島在2018對此建立的新法及影響 ⑤總結.

■ Prompt 2
☑ Response ①equality / Women's Day off / gender-related ②gender-balanced society ③Gender Equality Council / gender discrimination ④salary / pay gap, ⑤inequality / legally demand / work place equality ⑥fairer society.

■ Prompt 3
☑ Response ①but - 引入另一觀点 ②however - 儘管發生了某事，但仍起了改变 ③however 儘管有了進展，但仍未解決 ④50+ 引出原因 ⑤finally - 表達最終看法 / 作為 ⑥but - 引出影響及重要性.

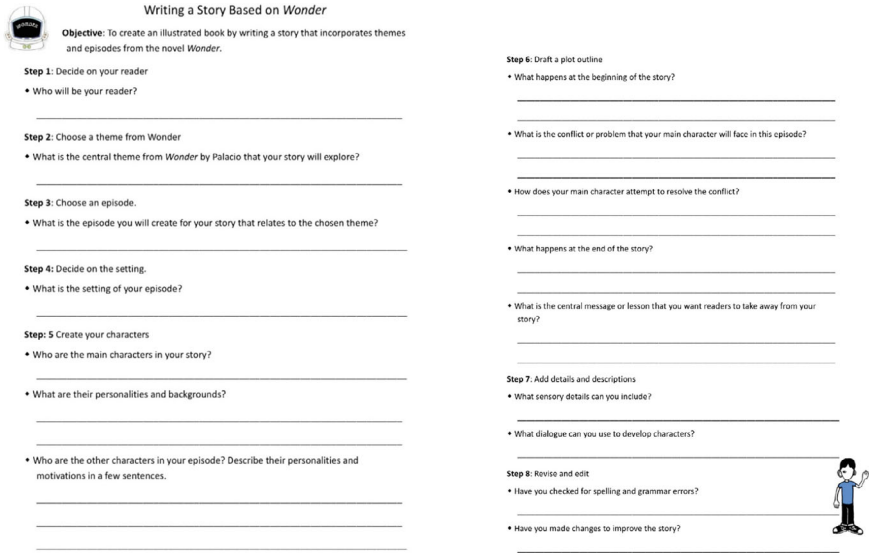
4. **Writing** (寫作)

■ Prompt 1 (Write down the types of errors pointed out by GPT that you made in the response)
☑ Response ①conflesh lighting → headlights on / ②right actrossing road → crossing the road / ③died → be lost / ④decease → prevent / ⑤animal-protecting → animal protection / ⑥and even → even though /

■ Prompt
☑ Response ①feel → experience / ②這3句例子開(and more 前)/ ③wildlifes habitat → wildlife habitats / ④will somtime be destroyed → are sometime destroyed / ⑤詞彙選擇 / ⑥稱寫問題 / ⑦物寫習選擇 / ⑧片語使用問題 / ⑨單詞問題.

→ My feedback: 雖然有人用 Chat-GPT 來製作作業或其他不正當用途，但我覺得它是個很好的提升英語基本功的工具，它能動選指出我的文法、字彙用法等等的錯誤，甚能夠對我的作文進行更進的修正，雖然比不上實際對話，但也沒有老師的指說下. Chat-GPT 是我最好的選擇. good!

Fig. 8.6 A student's completed worksheet recording the output from ChatGPT



Writing a Story Based on *Wonder*

Objective: To create an illustrated book by writing a story that incorporates themes and episodes from the novel *Wonder*.

Step 1: Decide on your reader

- Who will be your reader?

Step 2: Choose a theme from Wonder

- What is the central theme from *Wonder* by Palacio that your story will explore?

Step 3: Choose an episode.

- What is the episode you will create for your story that relates to the chosen theme?

Step 4: Decide on the setting.

- What is the setting of your episode?

Step: 5 Create your characters

- Who are the main characters in your story?
- What are their personalities and backgrounds?
- Who are the other characters in your episode? Describe their personalities and motivations in a few sentences.

Step 6: Draft a plot outline

- What happens at the beginning of the story?
- What is the conflict or problem that your main character will face in this episode?
- How does your main character attempt to resolve the conflict?
- What happens at the end of the story?
- What is the central message or lesson that you want readers to take away from your story?

Step 7: Add details and descriptions

- What sensory details can you include?
- What dialogue can you use to develop characters?

Step 8: Revise and edit

- Have you checked for spelling and grammar errors?
- Have you made changes to improve the story?

Fig. 8.7 A ChatGPT-generated worksheet that guides students in writing a story, dividing the writing process into eight steps, from deciding on the reader to revision and editing

Following the positive experience with ChatGPT, the teacher-researcher explored additional AI-powered educational tools, such as Diffit, Eduaide, Curipod, Twee, MagicSchool, and Gamma, with new tools emerging daily. These AI applications facilitated the rapid creation of lesson plans, slides, worksheets, and tests, often completing these tasks within minutes or even seconds. Take the AI tool Diffit, for example. In just a few minutes, it created a worksheet that includes a reading passage, comprehension questions, fill-in-the-blank vocabulary questions, and open-ended questions.

8.3.2 Semester Project for Elective Course "Read Wonder, Do Wonders"

The most notable aspect of the teacher-researcher's experience with AI occurred during a Grade 11 elective course that extended over an entire academic year, centered around R.J. Palacio's novel *Wonder*. During the first semester, students engaged in a thorough reading of the novel, including character and plot analysis, thematic exploration, and the examination of literary techniques. During the second semester, students embarked on a semester-long project. They were tasked with collaborating in teams to develop a website using Google Sites. The project required

the inclusion of several components: a succinct introduction to the characters, character images created with AI art tools, interactive avatars representing the characters, and illustrated storybooks.

After writing their own character introductions, students utilized ChatGPT to proofread their texts and documented both the original and revised versions. They then dedicated approximately two hours to generating character images using Bing Image Creator (renamed as Image Creator from Designer in January 2024) (Fig. 8.8). If needed, they could ask ChatGPT to modify their prompts. Feedback on the use of AI tools for image generation varied: one student reported that ChatGPT's prompt suggestions were not particularly effective, while another noted that the assignment provided significant learning opportunities (Fig. 8.9).

Meanwhile, students wrote their own stories inspired by *Wonder* and completed various worksheets, including a scaffolding outline worksheet (Fig. 8.7), the creative writing worksheet previously generated by ChatGPT, and a draft worksheet. The teacher-researcher then posted prompts on Google Classroom (Fig. 8.10), enabling students to request proofreading for their stories and record the ChatGPT-revised versions on their worksheets. Following the revision process, students created illustrations for their stories with either Bing Image Creator or Midjourney to generate images of characters from *Wonder* (Fig. 8.11). Finally, the students compiled their illustrated stories into formats such as flipbooks, audiobooks, or videos.

The most rewarding aspect of this project was witnessing students exceed expectations. The teacher-researcher was impressed by the websites created by her students using Google Sites. Each team's website featured concise introductions to eight designated characters from *Wonder*, videos of these characters' talking avatar, and the students' own illustrated books and/or videos. Students produced exceptional work, of which they can be justifiably proud. Examples of their accomplishments

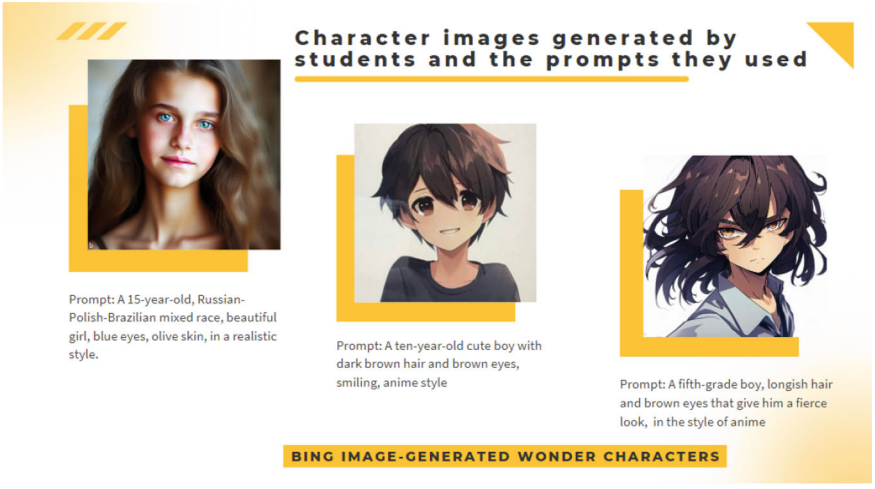


Fig. 8.8 Character images generated by students using Bing Image Creator

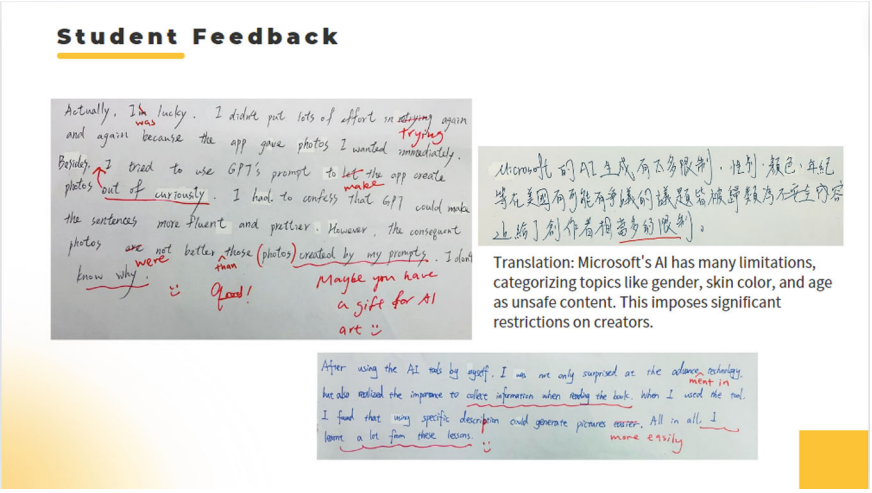


Fig. 8.9 Student feedback on using Bing Image Creator for image generation

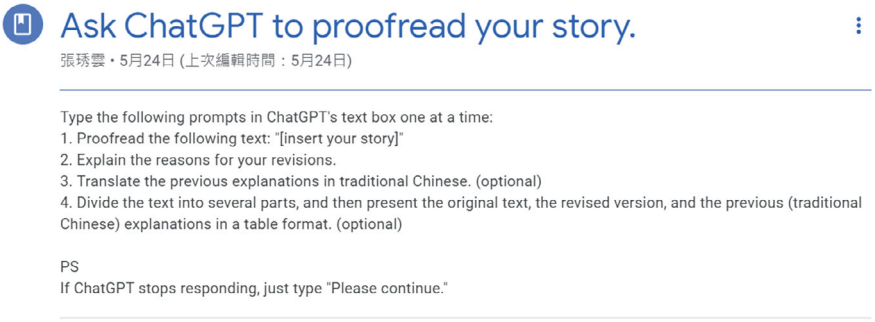


Fig. 8.10 Prompts posted on Google Classroom for students to request proofreading from ChatGPT for their stories

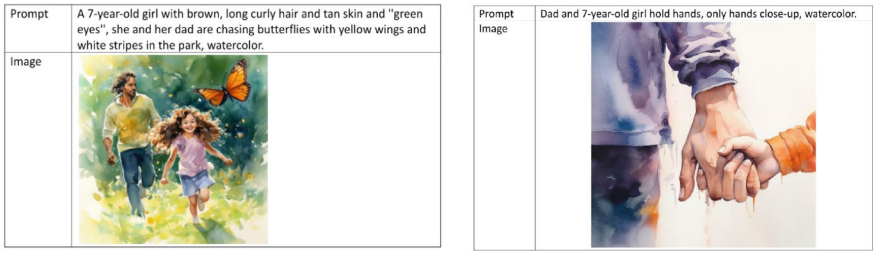


Fig. 8.11 Prompts devised by a student for Midjourney to generate images, along with the resulting illustrations

are presented below (Figs. 8.12 and 8.13). Notably, one student produced a remarkable video, incorporating voiceovers by her parents, which added a deeper emotional resonance to the project. Prior to students working on their illustrated books and videos, the teacher-researcher had used ChatGPT to generate a story. In comparison, the students' stories demonstrated greater emotional depth and sincerity.

These examples illustrate the teacher-researcher's use of AI tools in the classroom, demonstrating their impact on both efficiency and student creativity. The integration of these tools not only streamlined the author's workflow, significantly reducing the

Fig. 8.12 Two pages of a student's illustrated book

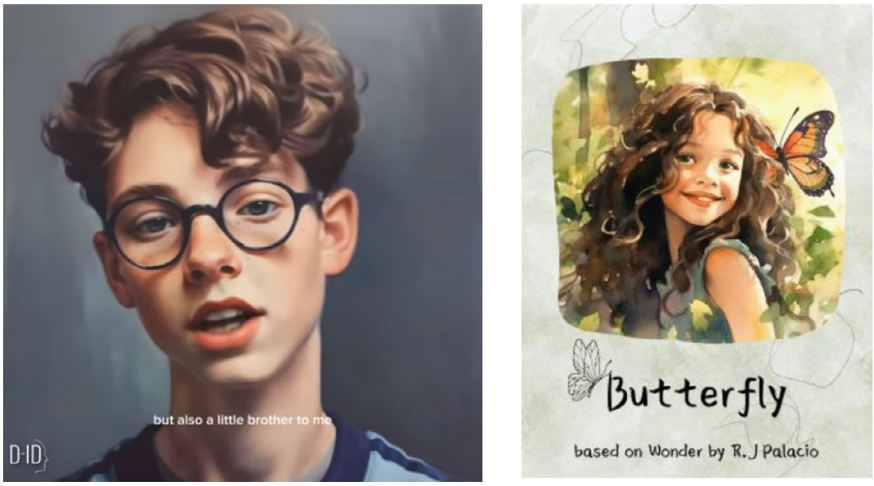


Fig. 8.13 (Left) A student's video of a character's talking avatar; (right) a student's video of her illustrated book

previously mentioned 50 weekly work hours, but also facilitated student engagement and creativity. Meanwhile, the teacher-researcher observed how students leveraged ChatGPT and AI art generators to produce innovative and emotionally resonant work. With ChatGPT's assistance, students with varying levels of English proficiency were able to create stories that were touching, humorous, or insightful. Similarly, AI art generators like Bing Image Creator enabled students with limited artistic skills to produce visually appealing illustrated books. The teacher-researcher is confident that these AI tools have expanded students' creative potential and encouraged them to exceed their own expectations.

According to McKinsey & Company, educational technology can enable teachers to allocate 20 to 30 percent of their time toward activities that support student learning (Bryant, 2020). Building on this, the teacher-researcher's experience with ChatGPT has proven to be even more time-saving. The efficiencies gained through these AI tools have allowed her to conceive innovative projects, such as the semester project mentioned above. However, the advent of AI tools raises a critical question: Do these technologies render teachers less necessary, or, to put it more positively, what aspects of teaching remain irreplaceable in the age of AI? It is imperative for educators to redefine their roles within the evolving landscape of language learning. As shown in the teacher-researcher's content, her students wouldn't have learned and created with AI's assistance without her guidance and creative lesson planning. Though AI tools can automate certain aspects of language instruction, they cannot replace the teacher's essential role in designing pedagogically sound activities and fostering critical engagement with technology.

Conversely, the integration of AI tools has enabled students to produce work that exceeds their typical capabilities, such as more polished writing and refined illustrations, providing a sense of achievement and motivation. Nevertheless, there is concern that some students may rely excessively on ChatGPT, potentially bypassing essential cognitive processes necessary for learning. As education specialist Momo Bertrand has remarked, "AI won't replace teachers—but a classroom revolution is coming" (Bertrand, 2023). While we have entered the era of artificial intelligence, it is crucial to consider how human intelligence and the unique skills of educators will continue to thrive and evolve alongside these technological advancements, ensuring a balanced and effective approach to language education.

8.4 ChatGPT, Teachers, and Their Students

As more and more people are using and developing AI-powered tools to facilitate reading and writing in English, questions such as "How can English language learners benefit from using AI-powered tools?", "How do English teachers use AI-powered tools to work efficiently?", or even "Will AI-powered tools replace English teachers?" have raised heated discussions. To investigate the real impact of AI-powered tools on English language teaching and learning, a mixed-methods study focusing on ChatGPT was conducted to answer three research questions:

1. What are the perceived advantages and disadvantages of using ChatGPT?
2. What factors contribute to the irreplaceability of human English teachers in class?
3. What factors predict students' optimal experience in using ChatGPT?

8.5 Methods

8.5.1 Participants

The participants were four experienced English teachers (2 females and 2 males) teaching senior high school students. On average, they had been teaching for more than 16 years. These teachers were purposefully selected because they were not only actively learning to use ChatGPT by themselves but also helping their peers to use this tool. In other words, they were helping fellow teachers to develop their technological pedagogical content knowledge and stimulate professional dialogue on the incorporation of AI into the TESOL profession. Their common motivation to use ChatGPT stemmed from a combination of curiosity, a desire to reduce workload, and increase work efficiency. They were self-taught learners, making use of various sources, including social media, open courses, online speeches, YouTube videos, and colleagues. They were pioneers of experimenting on using ChatGPT to teach and to work.

As well, two of the teacher participants agreed to invite their students to participate in this study. A total of 138 senior high school students (aged 16) successfully contributed their data.

8.5.2 Data Collection and Analysis

To answer the three research questions, data were collected from four teachers and the students of two of the teachers ($n = 138$). The data from the teachers consisted of (1) semi-structured online interviews on how they used ChatGPT at work (approximately two hours in total, conducted from late April to late May in 2023), (2) one classroom observation for 50 min, (3) several self-developed worksheets contributed by two participants, (4) online posts contributed by three participants, and (5) two online speeches given by two participants. For the qualitative data collected from the teacher participants, content analysis following the procedure suggested by Saldaña (2015) was adopted.

The data from the students consisted of their responses to three questionnaires measuring their motivational needs based on self-determination theory (Cheng et al., 2017; Ryan & Deci, 2013), optimal experience based on flow theory (Cheng et al., 2017; Csikszentmihalyi, 1997), and user experience (Wang et al., 2013). The questionnaires used in this study were adapted to fit the context. For motivation, an example item was: "I will proactively learn and apply knowledge related to using

ChatGPT to revise my English compositions". For optimal experience, a representative item was: "During the class just now, I was so immersed in the lesson that I lost track of time". Finally, regarding user experience, students responded to items such as: "ChatGPT helped me identify grammar mistakes that I hadn't noticed". For the quantitative data from the questionnaire responses, regression analysis was applied, using SPSS 21.

8.6 Results

8.6.1 What Are the Perceived Advantages and Disadvantages of Using ChatGPT?

Based on the interview data, as Table 8.1 shows, for teachers, using ChatGPT could lead to several advantages. Firstly, it enhanced their teaching proficiency. For example, they improved their English language skills, since they interacted with ChatGPT in English and they learned English expressions from the texts that ChatGPT generated. They also learned from ChatGPT how to provide constructive feedback to students, which they admitted was not a well-elaborated topic when they were receiving teacher education. Besides, they found it valuable in accessing quality learning materials, because they asked ChatGPT to generate examples and texts meeting their teaching goals. Second, ChatGPT improved their work efficiency by helping generate questions, examples, and exercises, and served as an inspiration for various creative learning activities. Moreover, teachers reported developing new or sharpening old skills. For example, they started being aware of the quality of the questions they asked and searching for evidence to confirm what ChatGPT had offered. They also engaged in altruistic sharing by exchanging knowledge and resources with their colleagues at school or fellow teachers online. These positive experiences led to positive emotions of confidence, relaxation, fun, and joy. As for the disadvantages, they only mentioned negative emotions, such as being worried about creating an image of "lazy teachers" and concerned about the potential loss of their professional skills, such as designing questions and learning activities.

From the perspective of teacher participants, there are advantages and disadvantages of students using ChatGPT. As Table 8.1 shows, students were found more engaged in class. They were attentive and showed a willingness to use ChatGPT and to complete tasks suggested by ChatGPT. Second, using ChatGPT resulted in high-quality production of student writing. Teachers could thus understand the ideas of their students instead of guessing from the context. Besides, with the support of ChatGPT, teachers offered quality supplementary learning materials, such as explanation and examples; this way, students could understand the content better. Next, ChatGPT assisted low achievers to generate English that conveyed their brilliant ideas. So, students had equal opportunities to contribute to the class. Additionally, ChatGPT served as a personal tutor, guiding students in their learning process. As a

Table 8.1 Advantages and disadvantages of using ChatGPT for teachers and students

Advantages	Disadvantages
For teachers	
• Teaching proficiency • Work efficiency • New & old skills • Altruistic sharing • Positive emotions	• Negative emotions
For students	
• Engaged • Quality production • Better comprehension • Everyone has a chance • Personal tutor • Building AI literacy • Positive emotions • Continuous using	• Overdependence • Laziness • Plagiarism • Misinformation • Feeling lost about the future

result, students did not have to wait for teachers' attention, while they were helping other students. In addition, it helped build students' literacy in AI technology. The more they used the tool, the more familiar they became with the tool. This experience could be transferred to the learning of other AI tools. Next, positive emotions such as satisfaction with their work and curiosity about what ChatGPT offered were prevalent among students. Finally, they also continued using the tool after class to solve problems in learning and life. However, there were also potential disadvantages for students, including overdependence on ChatGPT, tendencies toward laziness, plagiarism issues, exposure to misinformation, and concerns about their future path since they could be replaced by ChatGPT at the workplace.

8.6.2 What Factors Contribute to the Irreplaceability of Human English Teachers in Class?

For the second research question, four themes were extracted from the interview data (See Table 8.2). First, human teachers possessed a wealth of teacher knowledge, such as content knowledge, pedagogical knowledge, pedagogical content knowledge (Shulman, 1987), technological pedagogical content knowledge (Mishra & Koehler, 2006), etc. These types of knowledge enabled teachers to ask ChatGPT the right questions and evaluate the accuracy of the given information. Besides, with their professional training, they knew how to apply various strategies to scaffold the learning process, ensuring students' learning autonomy and progress (Ribbe & Bezanilla, 2013). Furthermore, teachers established interpersonal connections with their students, creating a supportive and engaging learning environment. They created and shared experiences in the same culture. Finally, human teachers leveraged various

Table 8.2 Four themes of research question extracted from the interview data

Theme	Significance
Teacher knowledge	Ask right questions Evaluate the accuracy of given information
Scaffolding strategies	Ensure quality educational experience
Interpersonal connection	Create supportive and engaging environment
Communication	Use verbal and non-verbal cues

means of communication, such as verbal and non-verbal cues, to facilitate effective learning experiences that go beyond the capabilities of AI systems.

8.6.3 What Factors Predict Students' Optimal Experience in Using ChatGPT?

To investigate factors influencing students' optimal experience with Gen-AI, multiple regression analysis was conducted with *Motivational needs* and *User experience* as predictor variables. The correlation matrix for these variables, presented in Table 8.3, shows significant positive correlations among all variables ($\gamma_{(44)} = 0.58 \sim 0.72$), indicating that higher levels of motivational needs and positive user experiences are associated with enhanced optimal experience.

The multiple linear regression analysis (see Table 8.4) reveals that the overall regression model for predicting optimal experience was statistically significant (Adj $R^2 = 0.38$, $F_{(2, 41)} = 13.90$, $p < 0.001$). Among the predictor variables, Motivational needs were found to be a significant predictor of Optimal experience ($\beta = 0.38$, $p = 0.035$), suggesting that students' intrinsic motivation plays a key role in enhancing their engagement and enjoyment when using Gen-AI for language learning. In contrast, User experience was not a significant predictor of Optimal experience ($\beta = 0.31$, $p = 0.085$), indicating that while user satisfaction is positively correlated with optimal experience, it does not independently drive the highest levels of engagement and enjoyment.

Table 8.3 Means, standard deviation, and pearson correlation matrix of motivational needs, user experience, and optimal experience

	M	*SD*	1	2
1. Motivational needs	6.41	0.66		
2. User experience	5.07	0.64	0.72^{***}	
3. Optimal experience	6.41	0.46	0.60^{***}	0.58^{***}

Note $N = 44$. *** $p < 0.001$

Table 8.4 Estimates of the impact of motivational needs and user experience on optimal experience

Optimal experience	Unstandardized Coefficient		Standardized Coefficient	
	B	*SE*	*β*	*t*
(Constant)	2.06	0.94		2.18*
Motivational needs	0.40	0.18	0.38	2.18*
User experience	0.44	0.25	0.31	1.76

Note. * $p < 0.05$

8.7 Discussion

The contribution of this study is three-fold. First, it presents teachers' first-hand experience and observation of using ChatGPT for their own preparatory work and in class with their students. The four teachers agreed that ChatGPT could be a valuable tool for teaching and learning English, with both teachers and students benefiting from using the tool to teach and learn. However, it is essential to address concerns about the potential negative impacts of using ChatGPT. So far, feasible solutions based on empirical evidence are still open for discussion. Second, the role that human English teachers play in the learning process is undeniable. Their capacities in various aspects exceed those of ChatGPT. Using ChatGPT should be considered as an addition to education instead of replacing teachers who have professional training and know how to build connections and communicate with students. Finally, the findings highlighted that student motivation, specifically the need for autonomy, competence, and relatedness (Ryan & Deci, 2002), significantly influenced their optimal experience when using AI tools like ChatGPT. Drawing upon Ryan and Deci's (2002) Self-Determination Theory, the results suggest that when students felt more independent (autonomy), capable (competence), and socially connected (relatedness), their experiences with AI technology became more engaging, satisfying, and meaningful. By fostering a supportive learning environment that meets these motivational needs, we can enhance the positive outcomes of integrating AI-powered language learning tools in English classes.

8.8 Conclusion

This chapter presents valuable experiences of integrating AI into TESOL, a synthesis of practical applications in authentic classrooms, and a small study with four teachers and questionnaire data from their students revealing the intricate relationship between technology and pedagogy. The findings illuminate not only the transformative potential of AI in enhancing teaching effectiveness and student engagement but also reveal complexities, including the roles teachers take, the emotional and cognitive demands they face, and the impact on students' learning experiences. Collectively, this chapter underscores how a thoughtful alignment of AI with pedagogical goals can lead to

substantial educational benefits, while also highlighting critical considerations for future implementations in TESOL classrooms. The collective experiences shared across the three sections weave a narrative of both promise and caution. Whether we like it or not, we have entered the era of AI.

As we contemplate the trajectory of TESOL in the age of AI, it is evident that a delicate balance must be struck. The potential benefits of AI, exemplified by efficiency gains and the unleashing of student creativity, must be tempered with a conscientious understanding of the nuanced drawbacks, such as overdependence and potential loss of intrinsic learning responsibilities. In crafting the future of language learning, it is not a question of AI replacing teachers but rather a call to redefine teachers' roles. The maximization of AI's potential in the field of language learning lies in its augmentation of human capabilities, serving as a tool that is used with consciousness and conscientiousness.

Appendix

Links to students' websites:

https://sites.google.com/view/nggyu-team-of-five
https://sites.google.com/smail.hc.edu.tw/wonder0621/home?authuser=0
https://sites.google.com/smail.hc.edu.tw/wonder-character-introduction/character?authuser=1
https://sites.google.com/smail.hc.edu.tw/wonder-boy/characters?authuser=0
https://sites.google.com/smail.hc.edu.tw/wondercksh/major?authuser=0
https://sites.google.com/smail.hc.edu.tw/wonderg6/home
https://sites.google.com/smail.hc.edu.tw/wonder/characters

References

Bertrand, M. (2023). *AI won't replace teachers—But a classroom revolution is coming.* Retrieved from Aljazeera.com: https://www.aljazeera.com/opinions/2023/5/24/ai-wont-replace-teachers-classroom-revolution-coming

Bryant, J., Heitz, C., Sanghvi, S., & Wagle, D. (2020). *How artificial intelligence will impact K–12 teachers.* Retrieved from Mckinsey.com: https://www.mckinsey.com/industries/education/our-insights/how-artificial-intelligence-will-impact-k-12-teachers

Buck, C., Heafield, K., & Ooyen, B. V. (2014). N-gram counts and language models from the common crawl. *Proceedings of the Language Resources and Evaluation Conference 2014,* 3579–3584. http://www.lrec-conf.org/proceedings/lrec2014/summaries/1097.html

Church, K., & Mercer, R. L. (1993). Introduction to the special issue on computational linguistics using large corpora. *Computational Linguistics, 19*(1), 1–24.

Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., & Fei-Fei, L. (2009). ImageNet: A large-scale hierarchical image database. *2009 IEEE Conference on Computer Vision and Pattern Recognition.* https://doi.org/10.1109/cvprw.2009.5206848

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, 4171–4186. Association for Computational Linguistics. https://doi.org/10.18653/v1/N19-1423&#8203

Meskó, B. (2023). Prompt engineering as an important emerging skill for medical professionals: Tutorial. *Journal of Medical Internet Research, 25*. https://doi.org/10.2196/50638

Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient estimation of word representations in vector space. *arXiv preprint*arXiv:1301.3781

Mishra, P., & Koehler, M. J. (2006). Technological pedagogical content knowledge: A framework for teacher knowledge. *Teachers College Record, 108*(6), 1017–1054. https://doi.org/10.1111/j.1467-9620.2006.00684.x

OpenAI. (2022). *ChatGPT* (Dec 20 version) [Large language model]. https://chat.openai.com

Ribbe, E., & Bezanilla, M. J. (2013). Scaffolding learner autonomy in online university courses. *Digital Education Review*, 98–112. https://doi.org/10.1344/DER.2013.24.98-112

Ryan, R. M., & Deci, E. L. (2002). Overview of self-determination theory: An organismic dialectical perspective. In E. L. Deci & R. M. Ryan (Eds.), *Handbook of self-determination research* (pp. 3–33). University of Rochester Press.

Shulman, L. S. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review, 57*(1), 1–22. https://doi.org/10.17763/haer.57.1.j463w79r56455411

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. *31st Conference on Neural Information Processing Systems* (pp. 1–11).

White, J., Fu, Q., Hays, S., Sandborn, M., Olea, C., Gilbert, H., Elnashar, A., Spencer-Smith, J., & Schmidt, D. C. (2023). *A prompt pattern catalog to enhance prompt engineering with ChatGPT*. arXiv:2302.11382. https://ar5iv.org/abs/2302.11382

Dr. Jo-Chi Hsiao is an Adjunct Assistant Professor at National Taiwan University of Science and Technology and an experienced English teacher at National Hsinchu Senior High School. Her research focuses on teacher professional development, TESOL, and Computer-Assisted Language Learning.

Shiowyun Chang obtained an MA in Foreign Languages and Literatures from National Taiwan University. She is an experienced English teacher at Hsinchu Chien Kung Senior High School and a seasoned translator, having translated 40 English books into traditional Chinese. In addition, she authored a translation guidebook for the GSAT test and co-authored a review and exercise book for the same test. Notably, in 2016, she guided students through the translation of More Nightlights into traditional Chinese, and the work was published the following year. In 2023, her teaching plan for an elective course titled "Read Wonder, Do Wonders" secured first place in a Ministry of Education-organized contest for digital learning.

Dr. Jason S. Chang received the PhD degree in computer science from New York University. He is a professor of computer science at National Tsing Hua University, Taiwan. His research interests span across the fields of natural language processing, computer-assisted language learning, information retrieval, and machine translation.

Chapter 9
Framework and Potential Applications of GAI-Based Language Learning Design

Yun-Fang Tu and Jing Chen

Abstract This study explores the framework and potential applications of Generative Artificial Intelligence (GAI) in language teaching and learning design, conceptualizing GAI as a tool, tutor, and tutee. From the language education and research perspective, we analyze the issues involved, with a particular focus on the core values of GAI and how it can enhance learner engagement. Firstly, the term GAI is explained, and its application is illustrated using real-life language learning contexts. The role of GAI as a tool, tutor, and tutee in the language learning process is discussed, and corresponding teaching strategies and examples are presented. In addition, GAI's critical influence and potential practical applications in language learning design are especially highlighted. For example, GAI has proven its excellence in content generation and learning path enhancement, providing learners with personalized learning experiences and effectively increasing learner engagement. In language learning contexts, GAI not only provides timely feedback as a learning tool enriched with a variety of listening, speaking, reading, and writing materials, but also acts as a tutor, providing personalized learning paths to meet the specific needs of students. In addition, students are taught to reflect on and consolidate their language knowledge and skills through AI, reflecting the role of GAI as a tutee. In conclusion, we discuss the research issues that may arise from using GAI in language learning. As GAI technologies advance and find application, the potential use of GAI technologies in language learning may bring more innovation and issues to the language learning field.

Keywords Generative Artificial Intelligence (GAI) · Language teaching · Language learning · Tools · Tutors · Tutees

Y.-F. Tu (✉)
Department of Data Science, Soochow University, Taipei, Taiwan
e-mail: sandy0692@gmail.com

J. Chen
Department of Educational Technology, Wenzhou University, Wenzhou, China

Y. Lan et al. (eds.), *AI-Mediated Language Education in the Metaverse Era*, Chinese Language Learning Sciences, https://doi.org/10.1007/978-981-95-0245-5_9

9.1 Introduction

With technological advancements, language learning approaches have undergone significant changes, especially with the rise of computer-assisted language learning (CALL) (Bin-Hady et al., 2023). Web-based Language Learning (WELL) and Mobile-Assisted Language Learning (MALL) have benefited from the widespread adoption of the Internet and mobile technologies, respectively (Beatty, 2013; Son, 2007). These technologies have not only made language learning more flexible and accessible, but have also enriched the approaches of teaching and learning (Parmaxi, 2023; Zhang & Zou, 2022). With the emergence of Artificial Intelligence (AI) technologies, the academic community has increasingly focused on Artificial Intelligence Assisted Language Learning (AIALL), especially Generative Artificial Intelligence (GAI). GAI-based learning language refers to the process of GAI technologies to support and enhance language learning. This approach typically involves using AI tools (e.g., chatbots) to automatically generate language learning materials, exercises, and interactions to provide a personalized and dynamic learning experience (Kohnke et al., 2023). For example, AI can simulate conversations, provide customized exercises, or adapt the content according to the learner's needs and progress, an approach that aims to leverage the highly adaptive and interactive nature of AI to enhance the efficiency and engagement of language learning (Jeon et al., 2023; Liu & Ma, 2023). Kohnke et al. (2023) studied the application of ChatGPT, a generative AI chatbot, in language teaching and learning, emphasizing the need for teachers and students to possess advanced digital skills to effectively use these tools and overcome their risks and limitations. They found that ChatGPT can simulate fundamental interactions, correct linguistic errors, create diverse texts, provide educational support, adjust conversational difficulty, and aid in multilingual learning. Additionally, Bin-Hady et al. (2023) analyzed data from 20 ResearchGate (RG) users over a two-week discussion, applying rootedness theory to explore ChatGPT's role in enhancing language skills (e.g., conversational practice, reading and writing abilities, grammar, vocabulary, and pronunciation) and its role in supporting language learning, including providing feedback, integrating the teacher's role, and assisting in language learning. These technologically rich learning environments not only provide significant opportunities for language learning, but also enable individuals to engage in personalized and self-directed learning anytime and anywhere, significantly increasing the convenience and efficiency of learning (García Botero et al., 2019; Hwang & Chen, 2023; Wang & Chen, 2020).

GAI technologies can automatically generate text, images, audio, and video content. Among these, text-generated and image-generated GAI have received particular attention (Lee et al., 2023). Some researchers have emphasized the importance of effectively integrating GAI into language teaching, which includes establishing educational objectives and teaching strategies (Cooper, 2023; Kohnke et al., 2023; Lai & Tu, 2023). Furthermore, Escalante et al. (2023) highlighted the key challenges faced by language education experts, especially in assessing and leveraging the potential of GAI in the language learning process. They encouraged experts to delve

into the areas where GAI currently excels in language instruction and to consider future trends in these domains. Additionally, they explored the unique advantages that human educators might maintain over AI in certain aspects. However, in the language learning field, GAI has been mainly focused on text generation, with other types of GAI needing to be further discussed (Jeon et al., 2023; Liu & Ma, 2023). Meanwhile, it is important to realize that GAI generally plays various roles in language learning.

Therefore, this study explored the key role of GAI in the design of language teaching, with a focus on analyzing the GAI framework and its potential applications in language instruction and learning. Additionally, this study aimed to explore potential research topics related to GAI-based language learning, with the goal of fostering academic advancement and practical innovation in this field.

9.2 Roles and Potential Applications of GAI-Based Learning Language

The application of AI in education is becoming increasingly widespread, and GAI-based language learning is a topic that has recently gained attention in the language learning field. These applications serve language learners and teachers, and cover automated written error correction, machine translation, conversational practice, generation of course content, presentation notes, and reading materials (Ingley & Pack, 2023). The quality of large-scale language models (LLMs) often exceeds the abilities of many people, including answering questions, recognizing errors, and creating content (Jeon & Lee, 2023; Lai & Tu, 2023; Pack & Maloney, 2023). Jeon et al. (2023) emphasized using multimodality in language learning, enhancing autonomous learning, deeper understanding, and extended learning experiences. Lan and Chen (2024) developed a collaborative team-teaching framework that integrates human teachers and GAI instructors, centered around three foundational concepts for constructing educational AI agents. These include:

1. A synergistic teaching approach, where human educators focus on interactive and personalized instruction, while GAI counterparts manage the control of the learning process and deliver instantaneous feedback;
2. The design of AI-driven pedagogical activities by human teachers, encompassing intelligent question-and-answer sessions and interactive learning modules;
3. The facilitation of student engagement with specially tailored AI instructional agents to support personalized and adaptive learning experiences.

Lan and Chen (2024) also delineated five critical design dimensions in this context: roles, procedures, rules/standards, guidance, and feedback mechanisms. Furthermore, GAI applications encompass multiple types, including text, images, audio, and video, demonstrating its capability to process various types of data and information (Hwang & Chen, 2023; Mondal et al., 2023). Text-GAI includes text generation and translation for semantic understanding; image-GAI involves classification

and labeling, which aids in visual data analysis and the creation of teaching materials; audio-GAI encompasses speech recognition and synthesis, which is essential in teaching and linguistic research; and video-GAI includes video analysis and generation, such as video content analysis, which is very helpful for creating instructional materials and media research.

Table 9.1 illustrates the diverse applications of Generative Artificial Intelligence (GAI) in language learning. Text-GAI focuses on text generation, enhancing the learning of listening, speaking, reading, and writing skills (Kumar & Kotler, 2024). Additionally, images can be generated through image-GAI to assist in teaching and learning, improving vocabulary memory and comprehension (Dastmalchi et al., 2024). Audio-GAI provides varied audio materials to enrich listening practice, while video-GAI integrates audiovisual materials, offering a comprehensive language learning experience. Moreover, the combined application of these GAI tools provides new methods and applications for language teaching and learning (Wang & Zou, 2024). The utilization of these different types of GAI offers new possibilities and opportunities for language instruction and learning.

From the perspective of educational applications, GAI assumes multiple roles in language learning. It facilitates a GAI-based learning language environment, enabling learners to engage in learning activities with GAI functioning as a tutor, tutee, and tool (refer to Fig. 9.1). Following the recommendations of Hwang et al. (2020) and Lai and Tu (2023), this study delineates GAI's roles into three categories: as a tool, a tutor, and a tutee.

9.2.1 GAI as a Tool

GAI serves as a learning tool for generating and providing language learning materials and resources, such as texts, dialogues, exercises, and so on, catering to various levels and styles of learning needs and enhancing learners' engagement and motivation (Ingley & Pack, 2023; Lai & Tu, 2023). As a learning tool, GAI aims to improve the efficiency and quality of learning and teaching while offering learners a richer, more diverse learning experience. This includes customizing or creating personalized exercises, tests, reading materials, and simulated dialogues and scenarios based on students' levels and interests. For instance, GAI could create customized exercises or mock exams based on students' learning history and preferences.

This study categorized the application of GAI in language learning in accordance with Bloom's Taxonomy, and Table 9.2 illustrates the role of GAI as a learning tool across various learning phases. As a learning tool, GAI plays a key role in language learning in various learning phases (Liu, 2024). In the memorization phase, GAI provides basic language learning materials, such as vocabulary lists, to assist students in memorizing fundamental vocabulary. In the understanding phase, it explains language rules and concepts, offering example sentences to facilitate comprehension. GAI offers simulated exercises and real-life application scenarios during the application phase, enabling students to apply their knowledge in practice. In the analysis

Table 9.1 Overview of the application of GAI in language learning

Types of GAI	Application of GAI in language learning	Description
Text-GAI	Teaching aspect: • Creates customized language exercises and tests that provide immediate feedback and interpretation • Develops lesson plans and materials adapted to different learners' needs Learning aspect: • Provides simulated conversations and writing exercises to enhance practical language application • Deepens understanding of language rules by explaining complex concepts and correcting errors	Listening: While Text-GAI primarily concentrates on text, textual descriptions can be utilized to augment the comprehension of listening content Speaking: Strengthens oral communication and responses by generating dialogue exercises and interactive quizzes Reading: Generates a variety of reading materials, such as stories and articles, to enhance text comprehension and analytical skills Writing: Offers guidance on different types of writing, including essays, reports, and creative writing, along with feedback and suggestions to improve writing skills
Image-GAI	Teaching aspect: • Utilizes images to assist in teaching new vocabulary and concepts to improve student comprehension and retention • Creates visual teaching materials to make the content more engaging and interactive Learning aspect: • Reinforces and deepens understanding of vocabulary and expressions through image-related activities • Uses images to stimulate thinking and inspire creative writing and oral expression	Listening: While there may not be a direct high relevance, combining audio content with images can assist learners in better understanding the listening materials Speaking: Enhances oral expression and vocabulary usage through practice describing the content of images Reading: Improves reading comprehension by utilizing images to aid in text interpretation Writing: Uses images as inspiration for creative writing tasks, such as describing pictures or composing stories based on images/comics

(continued)

Table 9.1 (continued)

Types of GAI	Application of GAI in language learning	Description
Audio-GAI	Teaching aspect: • Provides audio materials with different accents and pronunciation to enhance listening diversity; creates simulated listening tests to improve comprehension Learning aspect: • Engages in practice to recognize and comprehend diverse pronunciations; utilizes audio resources to simulate real-life language settings, improving practical language application abilities	Listening: Provides audio materials with various accents and intonations to improve listening comprehension and recognition skills Speaking: Provides simulated conversational practice to enhance pronunciation accuracy and speaking fluency Reading: Utilizes listening materials to assist in understanding text content, thereby improving reading comprehension skills Writing: Uses listening exercises to inspire writing, such as composing texts based on stories or conversations heard
Video-GAI	Teaching aspect: • Provides a blend of visual and auditory learning materials to aid students in gaining a more comprehensive understanding of the language and its associated culture • Develops contextual simulations using video to augment the interactivity and realism of the teaching experience Learning aspect: • Watching videos improves learners' comprehension of non-textual information, including speech, intonation, and body language • Utilizes videos to simulate real-life contexts, assisting learners in enhancing their ability to use the language in practical situations	Listening: Improves learners' listening comprehension of language by watching videos in different contexts Speaking: Practices oral expression and conversational skills by simulating the conversational scenarios presented in the videos Reading: Augments comprehension and analysis of textual content by integrating video content with subtitles Writing: Encourages learners to practice writing on topics relevant to the video content, such as describing scenes or creating stories related to the video

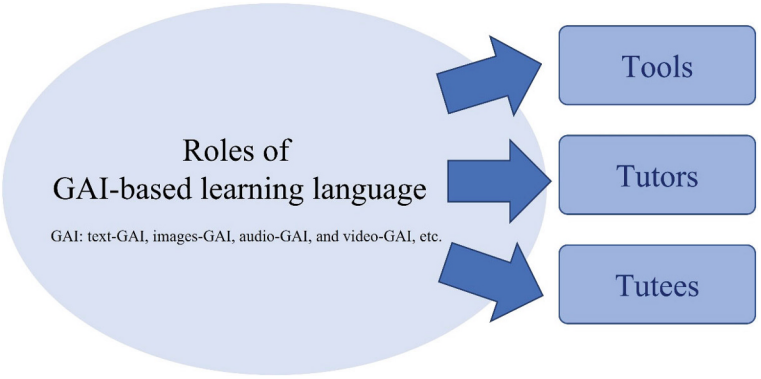


Fig. 9.1 Framework for GAI-based learning language

phase, GAI assists students in deeply analyzing language structure and usage. The evaluation phase provides assessment tools for students to gauge their language proficiency. Finally, in the creation phase, GAI supplies resources for creative language use, such as writing aids, to inspire students' creative expression.

For instance, in the study by Escalante et al. (2023), sample prompts sent to GPT-4 to generate feedback were as follows:

> "You will be a professional language teacher who is an expert on providing feedback on the writing of English language learners. Here is the writing prompt that students are given: [the weekly writing prompt was inserted here].
>
> Below I will share with you a student's writing. Based on their writing, comment on the following: …."

From Ingley and Pack's (2023) research on enhancing writing styles, here is an example prompt: You are an editor for a scientific journal. I have a draft of a/an (abstract/manuscript/proposal). Please review this draft and suggest where the writing style could be changed to conform with the academic writing style.

In addition, ChatGPT can be used as a tool to assist students in writing. An example of an English composition written by a ninth-grade student is as follows. First, this student copies the content to ChatGPT and asks him to reply in Chinese (i.e., I am a junior high school student, this is an English composition I wrote, please give me suggestions on what grammatical errors are in my composition according to the English level of a junior high school student, and answer them one by one in Chinese).

The process of student interaction and learning with ChatGPT is as follows:

Student:
In my study way I know a thing "the harder work, the happier I will be At first my English was very bad, at 7 grade my English score only 57, but this year my score turn well. Because my English tell me the great way to learn English. What's more I listen

Table 9.2 GAI as a tool in GAI-based language learning

Phase	Functions of GAI in language learning	Roles of the Teacher	Roles of the Student	Notes
Remembering	GAI provides repeated practice of vocabulary and grammar rules	Teachers select appropriate GAI materials and monitor students' learning progress	Students memorize foundational knowledge through repeated practice	It is ensured that the materials provided by GAI are suited to the students' learning level
Understanding	GAI explains language rules and provides example sentences	Teachers guide students in using GAI resources to ensure understanding	Students comprehend relevant language concepts and knowledge through GAI	At this stage, selecting GAI materials that are relevant and easy to understand is necessary
Applying	GAI provides simulated exercises and scenario applications	Teachers instruct students on how to utilize GAI for practical applications	Students apply their knowledge in simulated or real-life scenarios	Confirming and ensuring that GAI exercises align with actual application scenarios is essential at this stage
Analyzing	GAI assists students in analyzing language structures	Teachers assist students in selecting appropriate GAI analysis tools or teach the correct GAI analysis methods	Students use GAI tools to conduct in-depth analysis of listening, speaking, reading, and writing of language	At this stage, it is necessary to provide a variety of GAI analysis tools and in-depth examples to ensure the accuracy of the analysis and to develop students' GAI analysis literacy
Evaluating	GAI provides tools and methods for assessing language ability	Teachers validate and assist in interpreting assessment results and further providing or supplementing feedback	Students use the GAI tool to assess their language ability	At this stage, it is necessary to ensure the accuracy and adaptability of the GAI assessment tools
Creating	GAI provides services for generating creative language learning materials or works through programming prompts	Teachers foster students' creative thinking and provide instructional guidance	Students utilize GAI tools for creative expression	At this stage, the focus is on stimulating student creativity by providing GAI resources or collaborative services for creative works involving GAI

English songs from morning to evening such as Yesterday. This help me read book to early, then l read the books, I read book and listen the books, CD to learn how to speak English. Five months later My English was turn nice.
So From the learn experience, I know the harder I work, The happier I will be, I am a third-year junior high school student. This is an English composition I wrote. Please tell me what grammatical errors are in my composition according to the English level of a junior high school student, and answer them one by one in Chinese.
ChatGPT (as a Tool):

First of all, I would like to congratulate you on your remarkable progress in English learning. Your composition is already quite clear, but there are some grammatical errors and deficiencies in expression. Here are some guidance suggestions:
In the first sentence, "the harder work" should be changed to "the harder I work" because in this case, we need to use the subject pronoun "I" to connect the two clauses.
In the sentence "At first my English was very bad, at 7 grade my English score only 57, but this year my score turn well.", "at" should be replaced with "in", and "7 grade" should be changed to "grade 7". So the corrected sentence is: "At first, my English was very bad. In grade 7, my English score was only 57, but this year my score has improved."
In the sentence "This help me read book to early, then l read the books, I read book and listen the books, CD to learn how to speak English.", it should be corrected to "This helps me read books early, then I read books, I read books and listen to books on CD to learn how to speak English."
In the sentence "Five months later My English was turn nice.", it should be corrected to "Five months later, my English had become nice."
Overall, your English is quite good, but there are still some details you need to work on. Keep up the good work and your English will get better!

Student:

I want a more advanced expression, which words can I replace with which advanced vocabulary?

ChatGPT:

When you want to improve your expression, consider replacing some basic words with more specific and enriched ones. Here are some suggestions:
Replace 'the harder work' with 'the more diligent efforts'.
The original sentence "The harder I work, the happier I will be."
The revised sentence "The more diligent efforts I put in, the greater happiness I will attain."
Replace 'very bad' with 'extremely poor'.
The original sentence "At first, my English was very bad."
The revised sentence "Initially, my English was extremely poor."
Replace 'turn well' with 'showed significant improvement'
The original sentence "...but this year my score turn well."
The revised sentence "...but this year my score showed significant improvement"
Replace read book' with engage in reading literature'
The original sentence "This helps me read book to early, then I read the books..."

The revised sentence "This helps me engage in reading literature early, and then I delve into various books..."
Replace 'nice' with 'proficient'
The original sentence "Five months later My English was turn nice."
The revised sentence "Five months later, my English became proficient."

In general, try to use more specific and accurate vocabulary to improve your expression. At the same time, make sure that the words you use are appropriate to the context, and don't overdo the complexity to the detriment of your clarity of expression.

Here are some examples of learning activities using image-GAI.

Example 1: Picture Description and Story Creation

Purpose: To improve students' descriptive abilities and creative writing skills.

Activity Design:

- Generate a series of unique and imaginative pictures using image-GAI.
- Students describe scenes, characters, or activities in these pictures, practicing detailed descriptions and vocabulary in the target language.
- Then, students create a short story based on the pictures to practice storytelling and creative expression.

Example 2: Cultural Exploration and Comparison

Purpose: To deepen students' understanding of different cultural backgrounds.

Activity Design:

- Use image-GAI to generate pictures representing different countries and cultures, like landmarks, festivals, traditional clothing, etc.
- Students describe these pictures in the target language and explore each picture's cultural features.
- Discuss and compare similarities and differences between cultures to improve cross-cultural communication skills.

Another example of visualizing learning content is when a teacher, while preparing a course, uses image-GAI to assist in interpreting texts, thereby enhancing students' reading comprehension skills. Figure 9.2 shows the learning content for "My Home Town and My Country." The teacher extracts key points and uses GAI technology to generate an image of a London scene, which is then displayed during the learning activity. In the course summary phase, teachers can ask students to describe their favorite city, and then the students can also use GAI to create images to illustrate their favorite city based on their descriptions.

In the classroom, the teacher introduces "London's capital city, London, featuring Tower Bridge" to the students, and the teacher uses an image of "London's capital city, London, featuring Tower Bridge" generated through GAI to illustrate his fondness for London.

Cambridge, London and England

By Tony Smith

I come from Cambridge, a beautiful city in the east of England. It is on the River Cam and has a population of about 120,000. My home town is especially famous for its university. Many famous people studied here, such as Isaac Newton and Charles Darwin. There are lots of old buildings and churches to visit. Students and tourists enjoy trips along the river by boat.

Cambridge is 80 kilometers from London. London is in the south of England and it is on the River Thames. It has a population of about seven and a half million, so it is bigger and busier than Cambridge it is about 2.000 years old, and it is famous for Big Ben. Buckingham Palace and Tower Bridge. England itself is part of an island, and you are always near the sea. The small villages and beaches on the coast are popular for holidays. Tourists like the areas of low mountains and beautiful lakes in the north, and the hills and pretty villages in the south. Everywhere in England, you will notice how green the countryside is.

It is never very hot in summer or very cold in winter. So come and see England any time of the year, but bring an umbrella with you. You will need it most days.

Fig. 9.2 The learning content for "My Home Town and My Country."

In language teaching, using image-GAI for visualizing learning content is an attractive and practical method (Chen & Hu, 2024). Take the learning unit "A Story of Friendship" as an example, teachers can use this technology before class to transform text materials into comic form for use in classroom teaching. This method not only enriches the visual elements of teaching, enhancing students' in-depth understanding of the material, but also stimulates their creative thinking. In classroom activities, students are required to try to retell the dialogues of the characters in the comics in English, based on the comics provided by the teacher. This approach not only improves students' oral communication skills but also effectively promotes the development of their writing skills and creative expression.

In this framework, teachers are responsible for selecting and overseeing the use of GAI tools while providing guidance and feedback to enhance students' ability to effectively integrate GAI into their learning, thereby maximizing their educational outcomes (Zhang et al., 2024). Students, on the other hand, need to use GAI as an auxiliary tool to enhance their language skills. Importantly, GAI materials and tools need to be matched with the students' learning level and needs, ensuring the provision of high-quality content and feedback. In the future, GAI may include advanced memory techniques, interactive teaching, in-depth language analysis tools, as well as more personalized and comprehensive assessment tools to improve the efficiency and appeal of language learning.

9.2.2 GAI as a Tutor

GAI as a tutor is an important perspective of GAI-based learning language (Lai & Tu, 2023). Several researchers have noted that GAI offers automated language practice, grammar correction, and pronunciation instruction through personalized learning experiences and feedback, effectively enhancing language skills. It can tailor learning content to match learners' abilities and interests, and adjust teaching strategies to suit learners' current learning status. GAI exhibits high interactivity and simulates real conversations, providing learners with ample opportunities for language practice and increasing the appeal and practicality of learning. In other words, GAI can create a virtual teacher to provide immediate evaluation and advice on student assignments or spoken language. For example, a student submits a written or spoken paragraph, and GAI analyzes it, providing suggestions for improvement.

Table 9.3 displays the role of GAI as a tutor in different phases of learning. When GAI plays the tutor role in language learning, it provides support in different ways at different learning phases. In the memorization phase, GAI provides a guided service of repetitive practice of vocabulary and grammar rules to help students remember the essential elements of the language. In the understanding phase, language concepts are explained through examples to prompt students to explore and understand language knowledge. In the application phase, GAI creates authentic contextual simulations to guide students in applying what they have learned to real-world situations. The Analysis phase assists or guides students in analyzing language structures and gaining a deeper understanding of the structures of the language. In the Assessment phase, GAI evaluates students' language use and provides feedback services to assist them in adjusting and improving their language skills based on the feedback. Finally, in the creation phase, GAI guides students in creative language expression and inspires them to use their knowledge to express themselves innovatively. For example, the study by Kohnke et al. (2023) demonstrated a case where ChatGPT was used as a Tutor. In this instance, when teachers distribute texts to students, students may not fully understand certain vocabulary in the text. Moreover, learners can also request vocabulary explanations in their native language (such as Chinese) to facilitate better understanding.

In the study by Jeon and Lee (2023), one example of an instructional prompt was: "What is the difference between 'I am heading' and 'I am headed'?".

Another example involved role-playing: "I want to conduct a two-person role-play. This role-play is aimed at practicing the question-and-answer format, 'What are you doing?' and 'I'm …' Can you create some scripts for this role-play using only simple English?".

Another example of application provided is a reading comprehension question (see Fig. 9.3). When students come across unfamiliar vocabulary or complex sentences during the reading comprehension process, engaging in a dialogue with GAI is also a learning approach.

Students' results for ChatGPT based on Fig. 9.3 are as follows:

Table 9.3 GAI as tutors in GAI-based language learning

Phase	Functions of GAI in language learning	Roles of the Teacher	Roles of the Student	Notes
Remembering	GAI provides repetitive practice of vocabulary and grammar rules	Teachers monitor progress to ensure the material is appropriate for the students' level	Students memorize foundational knowledge through repetitive practice	In this phase, it is important to adjust the difficulty of the GAI exercises to the students' level
Understanding	GAI explains language concepts through examples	Teachers guide students to think, and ensure their understanding is accurate	Students explore and comprehend language knowledge	GAI explanations should be clear and combined with practical applications in this phase
Applying	GAI creates simulations of real-life situations for students	Teachers provide guidance and feedback to ensure correct application	Students apply their knowledge in the simulated scenario	In this phase, ensuring that the GAI simulation closely resembles the real-life environment is essential
Analyzing	GAI assists students in analyzing language structures	Teachers guide students using GAI for analysis, and provide supplemental learning materials	Students use GAI for in-depth analysis of language usage	In this phase, it is necessary to provide a variety of GAI analysis tools and examples to ensure the accuracy of the analysis, and to develop students' literacy in GAI analysis
Evaluating	GAI assesses students' language use and provides feedback	Teachers interpret and validate GAI's feedback, providing additional guidance and suggestions	Students make improvements based on feedback from GAI and teachers	In this phase, it is necessary to ensure that the GAI assessment is fair and comprehensive
Creating	GAI instructs students in creative verbal expression	Teachers encourage innovation, provide creative ideas, or stimulate student creativity	Students utilize their knowledge for creative expression	This phase stimulates creative thinking by working with GAI, and GAI can provide guidance

I was travelling on a British ship from India to London in 1905. Among all the passengers, no one would fail to notice Padishah, an Indian businessman. From the moment he got on the ship, he kept showing off the diamond on his turban. The big diamond successfully attracted attention, including that of the five ostriches on the deck .

One morning, as Padishah passed by the ostriches, one swallowed* his diamond suddenly. The bird ran around and mixed itself with the others in no time.

"Whose are these stupid ostriches? I want my diamond back!" Padishah shouted angrily. People began to crowd on the deck. I was among the first to arrive. I didn't feel half sorry for the man as he was asking everyone what to do for the diamond's return.

Word swept across the ship about the accident, reaching a man named Potter. After lunch in the hall, Potter went to Padishah and claimed, "I've contacted the ostriches' owner in London. He accepted my offer for all five and now I am the owner. I have the right to keep the birds and the diamond inside." I thought Potter was rather smart and regretted deeply at a lost chance.

Padishah was so upset and blurted, "There's no way you can get my diamond," he continued as he got calmer. "I can offer you 100 pounds for the birds." Potter simply refused and left.

After a while, Potter came back to the hall with a sly smile. He announced that he wanted to sell four birds separately to separate people by auction. The starting price was 80 for a bird. One of them, he would keep for himself.

Before the crowd realised what was happening, a diamond businessman got one bird for 100 pounds. The prices for the following birds rose higher and higher. After the fourth being sold, Padishah got half mad, not knowing what to do. Potter seemed suddenly sorry and said, "I should have kept them all. But... Trust me! The last bird must be the one."

I stared at the last bird for quite a while and asked Potter for a private conversation. After a long talk, I finally got the bird for 700 pounds._____

The next morning, I was waiting in line to get off the ship, picturing the amazing future life in mind. Then, I saw Padishah and Potter walk arm in arm on the shore, just like old friends. At that moment, I learnt a lesson worth a diamond.

27 . What happened to Padishah on the deck?
A . He was hurt by a passenger, B . He failed to get others' attention.
C . People shouted angrily at him . D . An ostrich took away his diamond.
28 . Potter came back to the hall to _____.
A . say sorry to Padishah B . tell people he was the new owner
C . sell ostriches at high prices D . talk with the diamond businessman
29 . Which of the following can he put into _____ in Paragraph 8?
A . How worried I was! B . What a great deal it was!
C . How upset Padishah was! D . What an honest man Potter was!
30 . What does the writer mean by the underlined sentence in the last paragraph? (Please answer in about 40 words)

Fig. 9.3 A question in a reading comprehension test

ChatGPT (as a tutor):
Certainly! I'd be happy to help you understand the word "blurted." In the context you provided, the word "blurted" is used in the sentence: "Padishah was so upset and blurted." In this case, "blurted" means that Padishah spoke suddenly and without thinking, perhaps in an impulsive or emotionally charged manner. When someone "blurts" something, it usually implies that they said something unexpectedly or without careful consideration.
So, in the passage, it suggests that Padishah, in his upset state, spoke out in a somewhat uncontrolled or spontaneous way, expressing his emotions without much thought. If you have any more questions or if there's anything else you'd like help with, feel free to ask!
Student:

Can you give me an example that further explains this vocabulary, with an example that is at the English level of a junior high school student?

ChatGPT:

Certainly! Imagine you're in a classroom, and the teacher is talking about an upcoming test. The teacher announces that the test will be more challenging than usual. Without thinking, one of your classmates raises their hand and says, "Oh no, that's not fair! Why are you making the test so hard all of a sudden?"

In this scenario, the classmate "blurted" out their reaction. It means they spoke suddenly and without thinking, expressing their frustration in a spontaneous manner. The classmate didn't plan what they were going to say; it just came out in the heat of the moment. "Blurted" is often used when someone says something impulsively or without careful consideration.

Student:

After the fourth being sold, Padishah got half mad, not knowing what to do. This is a sentence from the article, I don't quite understand the meaning of this sentence, can you explain?

ChatGPT:

Certainly! Let's break down the sentence: "After the fourth being sold, Padishah got half mad, not knowing what to do."
"After the fourth being sold" refers to the situation where four of the ostriches had been sold through the auction that Potter conducted. So, by this point, four of the ostriches were no longer available for Padishah.
"Padishah got half mad" indicates that Padishah became very upset or frustrated. The phrase "half mad" suggests a strong emotional reaction, but not necessarily complete madness. It emphasizes his intense feelings, possibly a mix of anger, frustration, and confusion. "Not knowing what to do" emphasizes Padishah's state of uncertainty. He was faced with a situation where he had lost control (with the auction of the ostriches) and was unsure about the next steps to take.
In summary, this sentence describes Padishah's emotional state after four ostriches were sold – he was highly upset and didn't know how to handle the situation. It adds to the narrative tension and sets the stage for the subsequent events in the story.

Student:

And what is the grammatical structure of the sentence?

ChatGPT:

The sentence "After the fourth being sold, Padishah got half mad, not knowing what to do." has a complex structure. Let's break it down:
Main Clause:
Subject: "Padishah"
Verb: "got"
Complement: "half mad, not knowing what to do"
Subordinate Clause (Beginning with "After"):"After the fourth being sold"
Subject: Implied (referring to the situation or the auction)
Verb: "being sold"
Object: "the fourth" (referring to the fourth ostrich)
Participial Phrase:
"being sold" functions as a present participle phrase that modifies the noun "fourth." It provides additional information about the situation, explaining what happened to the fourth ostrich.
The sentence uses a participial phrase ("After the fourth being sold") to provide additional information about the circumstances. The main clause then describes the

emotional state of Padishah and his uncertainty after the events mentioned in the subordinate clause. The use of participles and subordinate clauses adds complexity to the sentence structure.

Here are some examples of learning activities using image-GAI.

Example 1: Creating Images through Classical Poetry and Texts

Purpose: To enhance students' understanding and imagination of classical poetry and texts.

Activity Design:

- Use image-GAI to create images that match the scenes or emotions described in classical poetry and texts.
- Students are to observe these images and try to connect them with specific classical poetry and texts.
- The AI "teacher" will guide students in analyzing the imagery, metaphors, and symbols in the poetry and help them explain the meanings of the texts in their own words.

Example 2: Visualization of Stories about Idioms

Purpose: To assist students in understanding and memorizing idioms, improving their language application skills.

Activity Design:

- Use image-GAI to create illustrations or situational images for various idioms.
- Students need to guess the corresponding idioms based on the content of the images and explain their meanings and origins.
- The AI "teacher" will provide detailed explanations, examples, and scenarios of use of the idioms, helping students learn and apply them more effectively.

In this framework, teachers supervise and guide students, provide additional resources, and review and interpret GAI feedback to optimize learning outcomes. Students should actively participate, and utilize materials and feedback provided by GAI. It is important to note that the services provided by the GAI should be tailored to the learning needs of the students and to ensure that the content and feedback are of high quality. In the future, it may be possible to integrate more advanced technologies, such as virtual reality and augmented reality, as well as more personalized and dynamic learning paths to make the learning process more efficient and engaging.

9.2.3 GAI as a Tutees

In the role of GAI as a learning subject, the learner deepens his or her understanding of a specific language or skill by teaching it to GAI. The teacher still has an important role when GAI serves as a tutor. Teachers can use pedagogical techniques to design

units of study that assist students in gaining a deeper understanding of what they are learning. This approach is often realized through the concept that "teaching is the best way to learn," where learners reinforce their understanding and mastery by explaining and teaching new vocabulary, grammar rules, or other language concepts to GAI. This interactive approach encourages learners to think more profoundly and consolidate their knowledge. However, such research and teaching activities are uncommon (Tu & Hwang, 2023).

Table 9.4 illustrates the role of GAI as a tutee in different learning phases (López-Galisteo & Borrás-Gené, 2025). In this framework, students teach GAI to accomplish specific tasks. When GAI serves as a tutee in language learning, the student reinforces their learning by teaching GAI various linguistic elements. In the memorization phase, the student teaches GAI basic vocabulary and grammatical rules, and reinforces their memorization through this instruction. In the comprehension phase, students explain language concepts and rules to GAI and deepen their understanding of these concepts through instruction. In the Application stage, students teach GAI how to use language in specific contexts, applying what they have learned to real-world situations. In the analysis phase, students analyze GAI's use of language and make suggestions for improvement, thereby deepening their understanding of language structures. In the Evaluation phase, students assess GAI's linguistic comprehension and expressive skills, and adjust their teaching methods according to GAI's performance. Finally, in the creation phase, students use what they have learned to teach GAI innovative ways of expressing themselves.

An example is "Learning French: A case of student interaction with ChatGPT." In this learning scenario, we explore a student's interaction with ChatGPT aimed at enhancing their proficiency in French. The student adopts the role of a French language instructor, tasked with imparting knowledge of basic French grammar and vocabulary. ChatGPT, on the other hand, plays the role of a novice French learner, posing questions and absorbing information from the student's input.

The interactive learning process is as follows:

Student (as teacher):
"Today, we're going to learn about common greetings in French. First, 'Bonjour' means 'Good morning'. Try starting a conversation with 'Bonjour'."
ChatGPT (as learner):

"Bonjour! Comment ça va?"

Student:

"Great! 'Comment ça va?' means 'How are you?'. It's a good follow-up greeting. Now, try responding with: 'Ça va bien, merci. Et toi?'"

ChatGPT:

"Ça va bien, merci. Et toi?"

Table 9.4 GAI as a tutee in GAI-based language learning

Phase	Functions of GAI in language learning	Roles of the Teacher	Roles of the Student	Notes
Remembering	Students instruct GAI in basic vocabulary and grammar rules	Teachers oversee the students' teaching approaches and offer guidance	Students reinforce their acquired knowledge by teaching GAI	Ensuring that GAI accurately responds to the student's instruction is crucial in this phase. Additionally, it is important to be able to evaluate and provide feedback on the students' teaching methods
Understanding	Students explain language concepts and rules to GAI	Teachers guide students on how to teach effectively	Students deepen their understanding through teaching	In this phase, GAI can pose learning questions and provide learning feedback to ensure the students' understanding is accurate
Applying	Students teach GAI how to use language in specific contexts	The teacher provides real-life scenarios to guide student instruction	Students apply what they have learned in their instruction	In this phase, GAI can provide simulations of real contexts to test the effectiveness of the student's teaching. The correlation and accuracy of the actual and simulated scenarios should be carefully considered
Analyzing	Students analyze GAI's use of language and make suggestions for improvement	The teacher helps students understand complex language structures and guides them in transferring their knowledge to GAI	Students analyze GAI's performance to better understand language structures	In this phase, it is ensured that GAI can demonstrate different levels of language use

(continued)

Table 9.4 (continued)

Phase	Functions of GAI in language learning	Roles of the Teacher	Roles of the Student	Notes
Evaluating	Students can assess GAI comprehension and presentation skills	Teachers provide evaluation criteria and instruct students in how to conduct evaluations	Students adjust their teaching approach based on the performance of GAI, and reflect on their own learning	In this phase, it is necessary to ensure that GAI is able to model different levels of understanding so that students can assess them
Creating	GAI simulates the responses of natural learners to help students understand the effectiveness of instruction, and develop new ideas or creative works based on student instruction	Teachers are responsible for guiding students to explore innovative teaching methods that stimulate their creativity and innovative thinking	Students need to analyze GAI responses to learn from them and reflect on their knowledge and teaching methods	In this phase, it is necessary to validate and monitor GAI responses to ensure the accuracy and applicability of their feedback

Student:

"Perfect! Next, let's learn how to use these greetings in everyday conversations. I'll play a customer in a coffee shop, and you're the waiter. Remember to greet me with 'Bonjour'."

ChatGPT:

"Bonjour! Vous désirez?"
Student:"I'd like a cup of coffee. How would I say that in French?"

ChatGPT:

"You can say 'Je voudrais un café, s'il vous plaît'."

Student:

"Very good. You're learning quickly. Now, let's continue with other common French expressions."

Another example related to the learning of English grammar and adjective clauses involves students understanding and becoming proficient in the use of English adjective clauses through their dialogue with ChatGPT.

The process of student interaction and learning with ChatGPT is as follows:

Student (as a teacher):
I am a junior English teacher, now you are a junior student, today we are going to learn about the attributive clause, so what is the attributive clause? A "definite clause" in English is an "attributive clause" or a "relative clause". It is a subordinate clause that modifies a noun or pronoun and is used to further explain the antecedent. A definite clause is usually led by a relative pronoun or a relative adverb, including who, whom, whose, that, which, and a relative adverb, including when, where, why. Can you repeat the definition of the attributive clause?
ChatGPT (as a learner):

Certainly! An attributive clause, also known as a definite clause, in English is a subordinate clause that modifies a noun or pronoun. It is used to provide additional information and further explain the antecedent (the noun or pronoun it refers to). Typically, an attributive clause is introduced by a relative pronoun (such as who, whom, whose, that, which) or a relative adverb (including when, where, why). The main function of the attributive clause is to add descriptive or identifying details to the noun or pronoun it is modifying.

Student:

Very good, let's continue to learn, the relational words are divided into relational pronouns and relational adverbs, the relational pronouns are that, which, who, what, whom, those.
Relative adverbs are when, where, why.
For example: "This is the museum which was built last year". This sentence is a which-led attributive clause, do you understand?

ChatGPT:

Yes, I understand. In the example sentence "This is the museum which was built last year," the attributive clause is led by the relative pronoun "which." The relative pronoun "which" introduces additional information about the noun "museum" and connects the clause to its antecedent. The attributive clause "which was built last year" provides more details about the museum, specifying that it was constructed in the past year. This construction demonstrates the use of a relative pronoun in forming an attributive clause to modify and elaborate on a noun.

Student:

Perfect, then let us together to learn how to use these attributive clauses in daily life, I play the role of the customer, you play the salesman, remember to speak with me to use the attributive clauses.

Furthermore, within this framework, teachers are responsible for supervision, guidance, providing assessment criteria, and supporting innovative thinking, thereby assisting students in utilizing GAI as a learning tool. The role of students is to actively participate, using GAI as a teaching aid to consolidate and deepen their understanding of language knowledge. The focus is on the interactivity and adaptability between students and GAI, as well as the quality of feedback from GAI. Future applications may involve GAI simulating more complex learners, providing richer interactions

and challenges to enhance teaching and learning effectiveness (Fernández-Herrero, 2024).

In addition, as elucidated by Beck and Levine (2023), ChatGPT can function as a collaborative partner or a facilitative tool in the educational process. It not only assists students in mastering the intricacies of writing content, but also imparts essential writing skills. ChatGPT guides learners in exploring a range of writing styles and techniques, offering diverse drafting strategies, augmenting their writing prowess, and fostering creative thought processes. Furthermore, interaction with ChatGPT enables students to adeptly evaluate the revision impacts on various manuscripts, deepening their comprehension of the writing process and mitigating writing-associated anxiety. According to Jeon and Lee (2023), ChatGPT, as a GAI tool, assumes four pivotal roles in language education: interlocutor, content provider, teaching assistant, and evaluator. Functioning as an interlocutor, it facilitates role-playing and interactive language games, such as simulating a restaurant server or engaging in word-guessing games. As a content provider, GAI generates and recommends educational materials, such as grade-specific English role-play scripts and comic books, while also incorporating cultural insights. As a teaching assistant, it performs grammar checks, provides dictionary-style explanations, and activates prior knowledge. As an evaluator, ChatGPT delivers constructive feedback on student assignments, develops assessment tools, and analyzes conversational histories to assess student performance. Together, these roles highlight the versatility and adaptability of GAI in making language learning more interactive and efficient (Yeh, 2024).

9.3 Potential Research Topics Related to GAI-Based Language Learning

As mentioned above, GAI can play various roles in the language learning environment. Some scholars have pointed out the potential of GAI technologies in language teaching and learning, including rich input, interactive feedback, personalized learning, and language learners' motivation (Kohnke et al., 2023; Liu & Ma, 2023). For example, Zhou et al. (2023) conducted a comparative study to explore the differences in narrative theme writing between ChatGPT and Chinese Intermediate English (CIE) learners. The results showed that ChatGPT surpassed human writers in terms of narrativity, concreteness of words, and referential cohesion. However, compared to CIE learners, ChatGPT was found to need more syntactic simplicity and cohesion depth. Notably, after receiving update commands, ChatGPT improved syntactic simplicity but still significantly lagged behind CIE learners in the depth of cohesion. Additionally, like many existing technologies, GAI requires more careful consideration when integrated into language learning situations (Kohnke et al., 2023). For researchers in the field of technology-assisted language learning, considering how

to implement this in practice is both necessary and exciting. Therefore, many potential research themes have been raised for GAI-based language learning, including but not limited to the following:

- Integrating digital skills and GAI technologies in language learning and teaching. This research theme focuses on enhancing the digital skills of teachers and students so that they can efficiently utilize GAI technologies such as ChatGPT in language learning. Jeon and Lee's (2023) study emphasized the importance of teachers' professional knowledge when using GAI technologies in teaching. In addition, Beck and Levine (2023) emphasized the importance of teachers' professional guidance in assisting students in using GAI technologies critically and strategically, which can help students become broad-Bin-Hady minded and engaged writers.
- In addition, training programs for language teachers are also a key issue, which should include instructing teachers on how to select and utilize chatbots or other GAI technologies in language teaching and learning, how to customize or develop new chatbots according to the specific needs of their students, and how to assist students in developing the skills to interact effectively with GAI technologies in order to enhance their learning outcomes (Barrot, 2023; Cai et al., 2023; Hwang & Chen, 2023; Lan & Chen, 2024).
- Exploring learners' attitudes toward and behaviors regarding the use of GAI technologies to assist language learning, and their impact on higher-order thinking skills. This research direction focuses on how learners plan and adjust their learning behaviors based on their perceptions of using GAI technologies (e.g., ChatGPT) and how these attitudes and behaviors affect their higher-order thinking skills. In addition, some researchers have mentioned that teachers or instructional designers should pay special attention to the output quality of the GAI tools or systems as well as other factors (e.g., hedonic motivation) in designing the application of GAI technologies (e.g., ChatGPT) to assist language learning in enhancing positive attitudes toward ChatGPT among language learners (Cai et al., 2023). In addition, learners should decide whether to adopt ChatGPT to enhance language learning through rational assessment. This further highlights the importance of utilizing higher-order thinking skills when applying GAI technologies.
- Exploring the application and effectiveness of multimodal GAI technologies in language learning. Given that current research on GAI-assisted language learning primarily focuses on text-based GAI, it is recommended that the scope of integration of various GAI technologies (such as images, audio, and video) with language learning be included and evaluated. Its research focuses on how multimodal GAI technologies can enrich the language learning experience, improve language comprehension, and expression, and explore the application of these technologies to different language learning topics. Another area of interest is developing and assessing cross-modal learning strategies (which integrate text, images, audio, videos, etc.) in GAI-based learning language. Furthermore, it is crucial to analyze learners' engagement with GAI and their learning records. Such analyses provide educators with insights into the students' learning journey and

help identify challenges that might emerge while utilizing GAI technologies in educational contexts.

- Exploring the use of GAI technologies (e.g., ChatGPT) in the second language writing classroom and its impact on students' writing process. Explore how students use GAI technologies (e.g., ChatGPT) to assist in second language writing, particularly in the topic selection, initial research, outline creation, first draft writing, and self-assessment stages. For example, analyze the role of ChatGPT in these phases and how it affects students' creativity, critical thinking, and writing skills (Barrot, 2023). In addition, the study could further explore how students integrate ChatGPT output with reputable sources and personal experiences.
- Exploring the ethical issues and cultural bias of GAI technologies (e.g., ChatGPT) in language education. This issue focuses on researching how GAI technologies (such as ChatGPT) affect academic integrity and the potential for cheating in assignments and exams. It also explores the impact of its database and algorithms on students from different cultural backgrounds. Additionally, there is a need to establish effective mechanisms to ensure that the use of GAI technologies (like ChatGPT) adheres to educational and ethical standards and enhances its cross-cultural accuracy and inclusivity, especially in addressing the needs of non-English background students (Kohnke et al., 2023; Pack & Maloney, 2023; Yan, 2023). Some researchers have suggested creating guidelines and agreements for the ethical and responsible use of GAI technologies (such as ChatGPT) in language teaching (Barrot, 2023; Mohamed, 2023).

9.4 Conclusions

In conclusion, the successful integration of GAI technologies into language learning hinges on several pivotal elements, most importantly, the development of GAI literacy and higher-order thinking skills of educators and learners. Additionally, some researchers have pointed out that the key digital competencies required for effective use of GAI include: (1) technical proficiency, which involves understanding the functions and principles of GAI tools, effectively constructing prompts and interactions, and continuously updating related knowledge and skills; (2) pedagogical compatibility, which entails planning and implementing the use of GAI tools in language teaching and learning, as well as guiding learners to use GAI independently; and (3) social awareness, which means having a deep understanding of the limitations and risks of GAI tools, considering these factors when using them, and communicating to learners about the ethical and risky aspects of GAI (Hwang & Chen, 2023; Kohnke et al., 2023). Teachers must not only be adept at using GAI tools, but should also be capable of imparting this knowledge to students, fostering an environment where GAI is utilized effectively and responsibly. For students, adopting the correct mindset and attitude toward GAI is instrumental in leveraging its full potential to enhance their language learning experience. Additionally, understanding and

addressing the challenges and limitations of GAI plays a key role in its application in the teaching and learning process. By thoughtfully navigating these complexities, GAI could transform into a collaborative partner in the journey of language education, offering innovative approaches to teaching and learning, while preparing teachers and students for the future of AI in education.

Declaration of Generative AI in Scientific Writing During the preparation of this work, the author(s) used Grammarly and ChapGPT 4.0 to apply basic tools for checking grammar, spelling, etc. After using this tool/service, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the publication's content.

References

An, S., Zhang, S., Guo, T., Lu, S., Zhang, W., & Cai, Z. (2025). Impacts of generative AI on student teachers' task performance and collaborative knowledge construction process in mind mapping-based collaborative environment. *Computers & Education, 227*, Article 105227. https://doi.org/10.1016/j.compedu.2024.105227

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

Beatty, K. (2013). *Teaching and researching computer-assisted language learning* (2nd ed.). Routledge. https://doi.org/10.4324/9781315833774

Beck, S. W., & Levine, S. R. (2023). Backtalk: ChatGPT: A powerful technology tool for writing instruction. *Phi Delta Kappan, 105*(1), 66–67. https://doi.org/10.1177/00317217231197487

Bin-Hady, W. R. A., Al-Kadi, A., Hazaea, A., & Ali, J. K. M. (2023). Exploring the dimensions of ChatGPT in English language learning: A global perspective. *Library Hi Tech*. https://doi.org/10.1108/LHT-05-2023-0200

Cai, Q., Lin, Y., & Yu, Z. (2023). Factors Influencing Learner Attitudes Towards ChatGPT-Assisted Language Learning in Higher Education. *International Journal of Human-Computer Interaction*. https://doi.org/10.1080/10447318.2023.2261725

Chen, Y., Zhang, X., & Hu, L. (2024). A progressive prompt-based image-generative AI approach to promoting students' achievement and perceptions in learning ancient Chinese poetry. *Educational Technology & Society, 27*(2), 284–305.

Cooper, G. (2023). Examining science education in ChatGPT: An exploratory study of generative artificial intelligence. *Journal of Science Education and Technology*. https://doi.org/10.1007/s10956-023-10039-y

D'souza, N., & Dastmalchi, M. R. (2024, July). architectural creativity stranded at mid journey? Evaluating creative potential of prompts and images in generative AI. *In International Conference on-Design Computing and Cognition* (pp. 224–240). Springer Nature Switzerland.

Escalante, J., Pack, A., & Barrett, A. (2023). AI-generated feedback on writing: Insights into efficacy and ENL student preference. *International Journal of Educational Technology in Higher Education, 20*(1), 57. https://doi.org/10.1186/s41239-023-00425-2

Fernández-Herrero, J. (2024). Evaluating recent advances in affective intelligent tutoring systems: A scoping review of educational impacts and future prospects. *Education in Science, 14*, 839. https://doi.org/10.3390/educsci14080839

García Botero, G., Questier, F., & Zhu, C. (2019). Self-directed language learning in a mobile-assisted, out-of-class context: Do students walk the talk? *Computer Assisted Language Learning, 32*(1–2), 71–97. https://doi.org/10.1080/09588221.2018.1485707

Hwang, G. J., & Chen, N. S. (2023). Editorial position paper: Exploring the potential of generative artificial intelligence in education: Applications, challenges, and future research directions. *Educational Technology & Society, 26*(2). https://doi.org/10.30191/ETS.202304_26(2).0014

Hwang, G. J., Xie, H., Wah, B. W., & Gašević, D. (2020). Vision, challenges, roles and research issues of Artificial Intelligence in Education. *Computers and Education: Artificial Intelligence, 1*, Article 100001. https://doi.org/10.1016/j.caeai.2020.100001

Ingley, S. J., & Pack, A. (2023). Leveraging AI tools to develop the writer rather than the writing. *Trends in Ecology & Evolution.* https://doi.org/10.1016/j.tree.2023.05.007

Jeon, J., & Lee, S. (2023). Large language models in education: A focus on the complementary relationship between human teachers and ChatGPT. *Education and Information Technologies.* https://doi.org/10.1007/s10639-023-11834-1

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

Kohnke, L., Moorhouse, B. L., & Zou, D. (2023). ChatGPT for language teaching and learning. *RELC Journal, 54*(2). https://doi.org/10.1177/00336882231162868

Kumar, V., & Kotler, P. (2024). Transformative marketing with generative artificial intelligence. In *Transformative marketing: Combining new age technologies and human insights* (pp. 65–102). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-59637-7_4

Lai, C. L., & Tu, Y. F. (2023). Roles, strategies, and research issues of generative AI in the mobile learning era. *International Journal of Mobile Learning and Organisation.* https://doi.org/10.1504/IJMLO.2025.10059769

Lan, Y. J., & Chen, N. S. (2024). Teachers' agency in the era of LLM and generative AI: Designing pedagogical AI agents. *Educational Technology & Society, 27*(1), I–XVIII. https://doi.org/10.30191/ETS.202401_27(1).PP01

Lee, U., Han, A., Lee, J., Lee, E., Kim, J., Kim, H., & Lim, C. (2023). Prompt Aloud!: Incorporating image-generative AI into STEAM class with learning analytics using prompt data. *Education and Information Technologies.* https://doi.org/10.1007/s10639-023-12150-4

Liu, G., & Ma, C. (2023). Measuring EFL learners' use of ChatGPT in informal digital learning of English based on the technology acceptance model. *Innovation in Language Learning and Teaching.* https://doi.org/10.1080/17501229.2023.2240316

Liu, J. (2024). Enhancing English language education through big data analytics and generative AI. *Journal of Web Engineering, 23*(2), 227–249.

López-Galisteo, A. J., & Borrás-Gené, O. (2025). The creation and evaluation of an AI assistant (GPT) for educational experience design. *Information, 16*(2), 117. https://doi.org/10.3390/info16020117

Mohamed, A. M. (2023). Exploring the potential of an AI-based Chatbot (ChatGPT) in enhancing English as a foreign language (EFL) teaching: Perceptions of EFL faculty members. *Education and Information Technologies.* https://doi.org/10.1007/s10639-023-11917-z

Mondal, S., Das, S., & Vrana, V. G. (2023). How to bell the cat? A theoretical review of generative artificial intelligence towards digital disruption in all walks of life. *Technologies, 11*(2), 44. https://doi.org/10.3390/technologies11020044

Pack, A., & Maloney, J. (2023). Using generative artificial intelligence for language education research: Insights from using OpenAI's ChatGPT. *TESOL Quarterly.* https://doi.org/10.1002/tesq.3253

Parmaxi, A. (2023). Virtual reality in language learning: A systematic review and implications for research and practice. *Interactive Learning Environments, 31*(1), 172–184.

Son, J. B. (2007). Learner experiences in web-based language learning. *Computer Assisted Language Learning, 20*(1), 21–36.

Tu, Y. F., & Hwang, G. J. (2023). University students' conceptions of ChatGPT-supported learning: A drawing and epistemic network analysis. *Interactive Learning Environments.* https://doi.org/10.1080/10494820.2023.2286370

Wang, C., Wang, Y., & Zou, B. (2024). Revolutionising EFL pedagogy: Innovative strategies for integrating GAI (ChatGPT) into language teaching. *Journal of Language Teaching*, 4(1), 1–7. https://doi.org/10.54475/jlt.2024.004

Wang, H. C., & Chen, C. W. Y. (2020). Learning English from YouTubers: English L2 learners' self-regulated language learning on YouTube. *Innovation in Language Learning and Teaching, 14*(4), 333–346. https://doi.org/10.1080/17501229.2019.1607356

Yan, D. (2023). Impact of ChatGPT on learners in a L2 writing practicum: An exploratory investigation. *Education and Information Technologies.* https://doi.org/10.1007/s10639-023-11742-4

Yeh, H.-C. (2024). Revolutionizing language learning: Integrating generative AI for enhanced language proficiency. *Educational Technology & Society, 27*(3), 335–353. https://www.jstor.org/stable/48787034

Zhang, R., & Zou, D. (2022). Types, purposes, and effectiveness of state-of-the-art technologies for second and foreign language learning. *Computer Assisted Language Learning, 35*(4), 696–742. https://doi.org/10.1080/09588221.2020.1744666

Zhou, T., Cao, S., Zhou, S., Zhang, Y., & He, A. (2023). Chinese Intermediate English Learners outdid ChatGPT in deep cohesion: Evidence from English narrative writing. *arXiv preprint*arXiv:2303.11812. https://doi.org/10.48550/arXiv.2303.11812

Yun-Fang Tu is an Assistant Professor at the Department of Data Science, Soochow University, Taipei, Taiwan. Her research interests include GAI-based learning, information literacy, technology-enhanced language learning, artificial intelligence in education, mobile and ubiquitous learning, and bibliometric mapping analysis.

Jing Chen is a graduate student in the Department of Educational Technology at Wenzhou University, China. Her research interests include artificial intelligence in education, technology-supported subject education, and digital reading. She has five years of experience teaching English to junior high school students at a tutoring organization.

Chapter 10
Revolutionizing Language Learning: The Convergence of AI and the Metaverse in Educational Paradigms

Muhammad Saqlain Aslam and Yung-Hui Li

Abstract In the age of the metaverse, where virtual reality (VR) and mixed reality (MR) seamlessly blend with our daily lives, education is undergoing a remarkable transformation. Within this transformative landscape, Artificial Intelligence (AI) has demonstrated remarkable potential in various fields, including science and industry. In education, it presents a promising border for applying AI technology, where its transformative impact is increasingly evident. AI Advances enable tools for language learning and create more complex humans-physical world interactions. This chapter explores the intersection of AI and language learning within the metaverse era, revealing new frontiers in language education. While language acquisition has long been a dynamic field, integrating AI into the metaverse presents a groundbreaking opportunity. AI plays a pivotal role in these immersive spaces, from generating 3D objects and content without extensive human engineering to guiding the behavior of non-player characters (NPCs) for natural interactions. AI models like ChatGPT and GPT-4 empower virtual teachers and classmates to engage with students realistically. This enhancement provides lifelike interactions, breaking free from scripted responses and making the virtual classroom experience remarkably close to real-life human interaction. This chapter investigates AI's potential influence on education within the metaverse through a comprehensive literature review. We examine recent advancements, highlight key technologies and applications, and bridge the gaps between innovations and practical educational applications. The perspective given above comes from not only academia but also from industry, ensuring a wide spectrum of the practicality of AI is provided for further discussion.

Keywords Artificial intelligence · AI in Education · Generative AI · GPT · ChatGPT · Metaverse

M. S. Aslam · Y.-H. Li (✉)
AI Research Center, Hon Hai Research Institute, Taipei, Taiwan
e-mail: yunghui.li@foxconn.com

M. S. Aslam
e-mail: saqlain.msa@foxconn.com

Y. Lan et al. (eds.), *AI-Mediated Language Education in the Metaverse Era*, Chinese Language Learning Sciences, https://doi.org/10.1007/978-981-95-0245-5_10

10.1 Introduction

Over the past decade, the transformation of education systems has been significantly driven by technological advancements, with artificial intelligence (AI) playing a pivotal role. AI has not only emerged as a game-changer in various industries like healthcare, transportation, retail, and finance but has also gained prominence in the realm of education (Guan et al., 2020; Kamalov et al., 2023). In the educational sector, AI stands out as a particularly effective tool in language learning and teaching, potentially enhancing learners' academic outcomes. The application of AI in language education provides numerous advantages, revolutionizing students' learning experience and offering educators new teaching methodologies. In China, the government has significantly emphasized the significance of integrating AI technology into education, specifically focusing on its application in language learning. Furthermore, in addressing the teacher shortage issue, some scholars, such as Edwards and Cheok (2018), have even proposed substituting specific teaching roles with AI-driven robots.

Prominent technology giants like Google and Microsoft have been making substantial investments in AI-driven educational technology. Google's "AI for Social Good" initiative is dedicated to harnessing AI's potential to address significant global challenges, including those within the education sector. Microsoft has introduced a range of AI-powered educational tools, including Learning Tools for OneNote, which utilize machine learning (ML) to enhance reading and writing abilities (M. Liu, 2023).

As AI continues to reshape education, another technological concept has begun to intersect with it: the metaverse. The concept of an "education metaverse" has gained popularity worldwide. This idea envisions an immersive virtual environment where students engage in learning activities while interacting with their peers and AI-driven virtual instructors, creating a blend of physical and digital realities. AI technology plays a transformative role in this educational metaverse, particularly in language education. Many educators and researchers have started proposing future agendas and implementation scenarios in their learning practices. The increasing interest in the educational landscape may stem from a wide range of possibilities, including the virtual space that offers real-like representations of selves which possibly enhance the social aspect of teaching and learning.

This chapter examines the rapid growth of AI across various industries and its burgeoning importance in language education. It focuses on the advancements of AI and the metaverse within the educational domain, specifically in the context of AI-enhanced language learning. Additionally, the study discusses the opportunities, advantages, applications, key challenges, and risks associated with this innovative technology from an industrial perspective in language education.

The structure of our chapter is organized around four main sections. Section 2 offers a background overview of AI and the metaverse in the context of language education. Section 3 delves into the evaluation of AI's effectiveness in educational contexts. Section 4 explores the relationship between AI and the metaverse from

an educational perspective. Section 5 outlines future prospects and opportunities for using AI and metaverse in language education.

10.2 Literature Review

10.2.1 Artificial Intelligence in Education (AIEd)

The integration of AI into educational systems has been a transformative force, reshaping traditional paradigms of teaching and learning. This section of review explores the evolution of AI in education, its current state, and its particular significance in language learning contexts.

The term "Artificial Intelligence AI," first coined by John McCarthy in 1955, is defined as a computer's capability to perform tasks that typically require human cognitive skills, such as communication, reasoning, learning, and problem-solving. This definition, as Nilsson (1998), captures the essence of AI. Expanding on this, Baker and Smith (2019) describe AI as encompassing a broad spectrum of technologies and algorithms, including ML, natural language processing (NLP), data mining, and neural networks.

The journey of AIEd began in the 1970s, although its foundational technologies have developed since the 1960s. One of the earliest milestones was the development of ELIZA, an NLP program created by Weizenbaum at MIT (Weizenbaum, 1966). Eliza served as an early example of a machine-human interface and set the stage for future AI applications in education. Subsequently, Carbonell (1970) introduced SCHOLAR, a student-focused instructional program on South American geography, which provided instant feedback in natural language. This system began what Sleeman and Brown (1982) later termed Intelligent Tutoring Systems (ITS) in 1979.

Another significant development that contributed to ITS was the MYCIN project (Shortliffe et al., 1975). Designed to assist physicians in diagnosing and treating bacterial infections, MYCIN combined instructional methods with domain expertise, enhancing the accessibility of knowledge. The growth of AIEd has continued at a pace, with literature expanding to include a variety of themes and paradigms, reflecting the rapid advancements in AI technologies within the educational realm.

Despite the rapid advancement of AI technologies, their integration into educational practices has faced several challenges. Recent literature reviews have consistently highlighted critical issues in AIEd research, including a lack of educational perspectives, with many AIEd studies failing to incorporate robust educational theories and models, potentially limiting their practical applicability (X. Chen et al., 2020; Hinojo-Lucena et al., 2019; Zawacki-Richter et al., 2019). There is also a significant gap between the theoretical capabilities of AIEd technologies and their actual implementation in authentic educational settings (Bates et al., 2020; Kabudi et al., 2021). Furthermore, AI-enabled e-learning research over the past two decades has

shown a concerning lack of grounding in educational theories and models (Tang et al., 2023). These challenges underscore the need for more collaborative efforts between AI developers and educational institutions to bridge the gap between technological innovation and pedagogical implementation.

10.2.2 AI in Language Education

While AIEd has been extensively studied in general education contexts, its specific application in language education remains a relatively unexplored frontier. The growing interest in this area is evident from the increasing number of publications on AI-enhanced language learning between 2000 and 2019 (Fig. 10.1). However, most existing studies have primarily focused on examining specific AI tools and applications in language learning contexts. There is limited research exploring broader AI trends and their potential to transform language education. This gap in literature presents both a challenge and an opportunity for researchers and practitioners in the field.

Given the current state of AIEd and its emerging role in language education, this chapter aims to advance AI technologies in the context of language education, making meaningful contributions to the broader language education community. It seeks to deepen our understanding of AI in language education, encompassing its present status, potential benefits, challenges, and anticipated future developments.

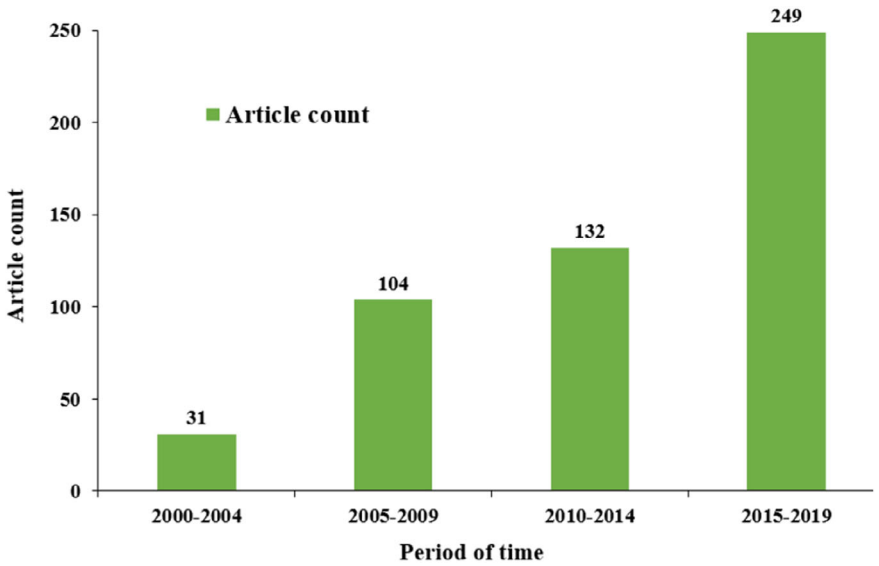


Fig. 10.1 Frequency of articles on AI-enhanced language learning (2000–2019) (X. Huang et al., 2023, Fig. 2)

The study also aims to bridge the gap between theoretical AI capabilities and practical implementation in language learning environments while providing insights for educators, academic researchers, AI technology developers, and other stakeholders in the language education ecosystem. By addressing these objectives, we hope to contribute to the evolving landscape of AI-enhanced language learning and pave the way for more effective, theoretically grounded, and practically applicable AIEd solutions in language education.

10.2.3 Metaverse in Education

The term "Metaverse" is formed by combining the prefix "meta," signifying transcendence or going beyond, with the word "universe," which denotes a parallel or virtual environment connected to the physical world. While not entirely new, the concept gained widespread attention after Mark Zuckerberg officially announced the Metaverse project in October 2021.

The metaverse has roots in science fiction literature, as seen in works like "Snow Crash" (Stephenson, 1994). In Snow Crash, the characters transform into avatars and operate within a three-dimensional (3D) virtual reality environment known as the metaverse. It also gained attention through the movie adaptation of the novel, titled "Ready Player One" (Cline, 2011). Prior to the current surge in interest, platforms like Second Life and World of Warcraft had already captured millions of users, demonstrating the appeal of virtual worlds (Wiederhold, 2022).

In education, the concept of the metaverse is not novel, as numerous researchers and educators have explored its potential impact on language learning (Fig. 10.2). For example, a study conducted by Kemp and Livingstone (2006) delved into integrating the metaverse using a virtual world called "Second Life" with learning management systems, aiming to improve the learning experience. Collins (2008) proposed the metaverse as a future space for social interactions in higher education. Schlemmer (2014) argued that the 3D digital virtual world with avatars enhances interaction and communication. More recently, Hwang and Chien (2022) explored the functions of intelligent tutors, tutees, and peers in delivering educational services in the metaverse from an AI perspective.

Despite this growing body of work, the metaverse remains a relatively unexplored concept in language education. This chapter evaluates the current state of research on this topic and provides a comprehensive overview of its potential applications and challenges in educational contexts.

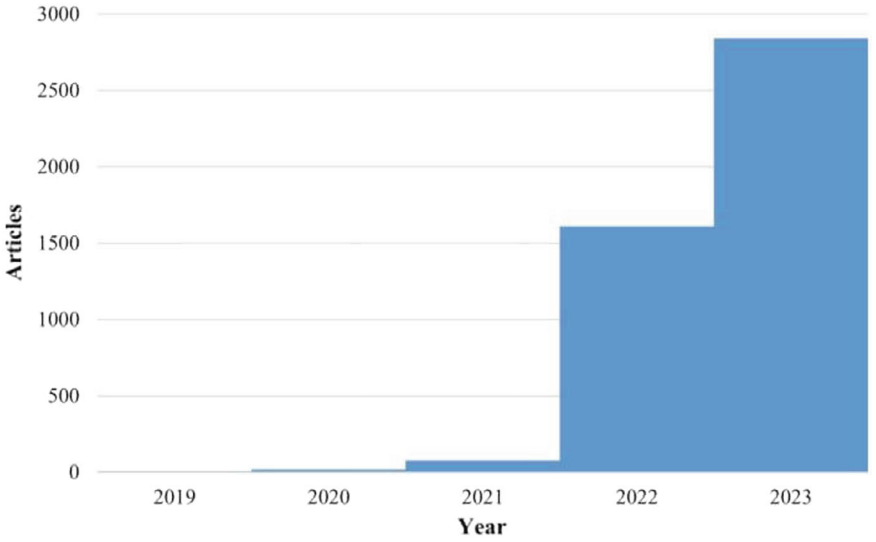


Fig. 10.2 Metaverse-related publication growth (Abilkaiyrkyzy et al., 2023, Fig. 1)

10.2.4 Generative AI in Language Learning and Metaverse Education

Recent advancements in ML have given rise to generative AI, a cutting-edge technology with significant potential to enhance language learning processes (Grassini, 2023). Generative AI creates original content across various media formats, including text, audio, and interactive scenarios, by analyzing and learning from patterns in training data (Andriulli et al., 2022; Lim et al., 2023). In the context of language education, academic research has explored diverse applications of generative AI. For instance, Mageira et al. (2022) examined the impact of AI chatbots on content and language-integrated learning, demonstrating the potential for immersive language practice in virtual environments. Bowles and Kruger (2023) investigated the influence of generative AI technologies, including ChatGPT, on college writing, which has implications for language skill development in metaverse-based educational settings.

While not explicitly focused on language learning, studies by Huh (2023) on ChatGPT's capabilities in medical education and Qadir (2023) on ethical considerations in AI implementation provide valuable insights for the integration of generative AI in immersive language learning environments. As the convergence of AI and metaverse technologies continues to evolve, these findings contribute to our understanding of how generative AI revolutionizes language education through personalized, immersive, and interactive learning experiences.

10.2.5 Generative AI in Multilingual Translation Technology

Generative AI technology promotes a diverse and inclusive experience by facilitating communication in various languages. Razumovskaia et al. (2022) elucidated multilingual translation technology functions by employing NLP algorithms to identify the user's input language and then translate it into the language of other users. This technology serves various purposes, including enabling users to converse in multiple languages and providing automatic translation for text-based interactions. Additionally, it supports automatic translation for audio conversations, facilitating communication among users who may not share a common language.

As per findings from Natarajan et al. (2022), multilingual translation technology offers numerous advantages. It substantially diminishes the time and expenses of manual translation, enabling businesses to introduce their products and services to international markets swiftly. Moreover, it delivers precise translations between languages, facilitating mutual comprehension among individuals from diverse cultural backgrounds. Y. Liu et al. (2020) expounded on the emergence of generative AI-based multilingual translation technology, which automates translations across various languages, potentially revolutionizing language learning methods.

Multilingual translation technology has found extensive applications in language education, international travel, and content sharing on social media platforms, granting learners access to content in multiple languages (Lv, 2023). Figure 10.3 provides an example of a comprehensive content creation application utilizing generative AI, which could be adapted for educational purposes.

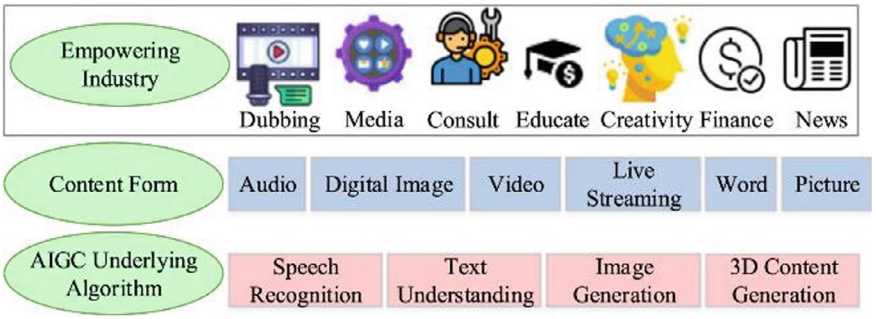


Fig. 10.3 A unified content generation application powered by generative AI (Lv, 2023, Fig. 3)

10.3 The Evolution of AI Technologies in Language Learning

In today's fast-paced world, technological integration has revolutionized diverse facets of education and one of the most impactful changes is evident in language learning. Language researchers have increasingly focused their attention on the integration of AI-powered technologies into language learning and teaching. This adoption of AI in language education holds the promise of improving the learning experience for students through the facilitation of personalized, interactive, and communicative learning processes.

The education landscape has experienced profound transformations with the prevalent adoption of digital technologies and the expansion of interconnected information networks (Khang et al., 2023; Wei, 2023). Language educators have embraced AI to develop virtual language learning environments that immerse learners in dynamic experiences, effectively streamlining the language learning process. These developments have brought forth a range of advanced AI tools that facilitate complex activities essential to the educational process (Fullan, 2023; Wright et al., 2023).

As educational models evolve to accommodate the changing landscape of AI-driven learning technologies, the role of AI in enhancing teaching and learning processes has grown significantly. Over the past decade, a paramount concern has been the effective utilization of AI technology to improve language teaching and learning (Ekin et al., 2023). In response to this concern, many AI-powered learning management systems (LMS) have been suggested by Alfalah (2023) to improve the educational experience in multiple ways.

Researchers emphasize that while AI learning technologies offer significant enhancements to both language teaching and learning, they also introduce a variety of challenges related to interactions between students and these diverse technological tools, given their various forms and sizes (Anderson, 2023). The effectiveness of students' language learning experiences and outcomes through modern AI-driven learning aids is notably influenced by factors such as their attitudes, emotions, and the quality of their learning experiences (Elme et al., 2022).

10.3.1 Pedagogical Implications of AI in Language Education

The ongoing evolution of AI has led to significant breakthroughs, substantially influencing language teaching and learning practices (X. Chen et al., 2020; Guan et al., 2020; Kamalov et al., 2023; Wei, 2023). Today, AI innovations in language education have moved beyond theoretical concepts in laboratory settings to practical applications in complex, real-world learning environments.

In language education, AI technologies possess substantial potential to revolutionize conventional instructional approaches by delivering personalized learning

experiences customized to individual demands and preferences. Companies have introduced innovative solutions to language learning, including Individual Adaptive Learning Systems for personalized learning, Aided Teaching Systems to enhance classroom management, grading, and evaluations, and systems to address second-language challenges.

Integrating AI into language education systems is not only reshaping the dynamics of learning, teaching, and institutional operations but also marking a substantial shift in the future of language education, where AI's data analysis capabilities identify effective teaching methods and adapt content accordingly. In line with this educational direction, educators have seamlessly integrated AI-assisted language learning tools into their teaching methods to assist learners in improving their language skills.

10.3.2 AI-Assisted Language Learning Tools

AI is an intelligent system capable of accomplishing diverse language-related tasks (Aldosari, 2020). For instance, individuals utilize AI-powered tools to seek answers to language-related questions, and these tools swiftly deliver the necessary information. AI applications in language education settings make intelligent decisions like human decision-making. Moreover, AI is extensively employed in language learning to enhance learners' language skills and subskills (Xia et al., 2023; R. Zhang & Zou, 2022).

Numerous AI-assisted language learning tools are now accessible on both computers and mobile devices, greatly enhancing language learning accessibility. These tools offer valuable support for improving a wide range of language skills. Among them, ChatGPT, a prominent AI-powered language learning tool, distinguishes itself by its potential to benefit learners across various language skills and subskills (Baskara, 2023; Hong, 2023; Kohnke et al., 2023). ChatGPT aids learners by providing writing ideas, suggesting alternative sentences to enhance writing performance, and facilitating significant progress in their language learning journey (Su et al., 2023; Yan, 2023). Moreover, ChatGPT understands human queries and provides optimal responses. It is worth noting that AI-supported language learning tools, including ChatGPT, are known for creating immersive and engaging learning environments, making language learning tasks convenient, and ultimately boosting overall language proficiency (Divekar et al., 2022).

10.4 Real-World AI Applications in Language Education

The integration of AI technology in language education has shown great promise in delivering tailored learning experiences, enabling dynamic assessments, and fostering meaningful interactions within online, mobile, or blended learning environments (K. Zhang & Aslan, 2021).

This section explores key AI applications that are transforming language learning:

10.4.1 Natural Language Processing (NLP)

NLP empowers machines to comprehend human language and is a valuable language learning tool. It provides the functionality of machine translation (MT), enabling the automatic conversion of a source language into a target language. NLP technology offers personalized learning materials by analyzing a student's language performance and generating customized content to address their specific needs, such as vocabulary lists for struggling students. NLP technology enhances language learning by analyzing authentic language data, helping students apply language skills in real-world contexts, and making learning engaging and relevant (M. Liu, 2023).

10.4.2 Chatbots for Language Practice

A chatbot, also referred to as a bot, chatterbot, dialogue system, conversational agent, virtual assistant, or virtual agent, is a software application that engages users through chat, simulating human conversations by responding to questions and prompts using text or audio. Chatbots have gained significant traction in language education, offering learners opportunities for authentic communication practice.

Recent research has demonstrated their efficacy in various aspects of language acquisition (Wollny et al., 2021). Researchers have increasingly focused on elucidating how educational chatbots enhance learning across diverse contexts, including language learning (Kuhail et al., 2023), vocabulary acquisition (T.-H. Huang & Wang, 2023; Yunjiu et al., 2022), and developing communicative competence (Hu & Hu, 2020; Kim et al., 2022).

By simulating conversations and providing immediate feedback, chatbots motivate learners to actively engage with the target language, positively impacting their learning journey (Okonkwo & Ade-Ibijola, 2021).

10.4.3 Intelligent Tutoring Systems (ITS) for Language Instruction

An ITS is a computer program that offers personalized instruction or feedback to students, typically without direct involvement from a human instructor. ITS employs AI algorithms to provide personalized language instruction and feedback, offering tailored guidance on grammar, vocabulary, and pronunciation. Thus, it bridges the gap between classroom instruction and individual tutoring (Guan et al., 2020; Kamalov

et al., 2023). The adaptive nature of ITS makes them particularly valuable for addressing the diverse needs of language learners.

10.4.4 AI-Driven Personalized Language Learning

AI-driven personalized learning involves tailoring educational materials and experiences to accommodate each student's unique requirements, learning preferences, and styles. This approach empowers students to advance at their own pace, resulting in increased engagement and improved learning outcomes (K. Zhang & Aslan, 2021).

A study by Shannon and Chapelle (2017) discussed how AI-driven adaptive learning systems provide customized grammar exercises and reading materials based on a learner's current level and progress. Furthermore, a study conducted with high school students learning Spanish (Godwin-Jones, 2017) revealed that connecting language tasks to students' personal interests outside of school could enhance motivation and retention within an ITS.

These findings underscore the potential of AI-driven personalization to significantly improve various aspects of the language acquisition process, from vocabulary and grammar to reading comprehension and conversational skills.

10.4.5 Virtual Reality (VR)

VR, or Virtual Reality, is a computer application that immerses users in a 3D environment through visual and auditory simulations. It is known for its ability to enable users to explore virtual surroundings and interact with virtual objects (Guan et al., 2020).

Integrating AI and VR technology has proven effective in immersive language learning and improved reading comprehension among young learners (Ijaz et al., 2017). A study involving an intelligent glass system confirmed the beneficial role of AI technology with visualizations as a social communication aid for children and adults with autism (Keshav et al., 2017). While promising, the implementation of VR in language learning also presents challenges, such as potential cognitive overload. This highlights the need for careful design and further research into AI-enhanced VR applications for language education. Further research and development in these areas will be crucial to fully harness their capabilities and address the complex needs of language learners in the digital age.

10.5 AI Advantages in Language Learning

The integration of AI technology is revolutionizing language education, offering unprecedented opportunities for enhanced learning outcomes and global access to quality language learning resources. This section explores the transformative impact of these technologies on language acquisition and practice.

10.5.1 Enhanced Learning Outcomes

AI-driven language learning platforms are at the forefront of personalizing the educational experience. These systems evaluate a learner's proficiency level, learning style, and pace, customizing lessons to provide appropriately challenging content. This tailored approach not only makes the learning process more efficient but also significantly more engaging.

Immediate feedback is crucial for students to assess their progress and enhance their learning, a feature often lacking in traditional language learning methods. AI effectively addresses this issue by delivering real-time feedback on pronunciation, grammar, and vocabulary usage through AI-powered language apps and tools. This support empowers learners to practice speaking and writing with confidence, as they know they will receive instant corrections and helpful suggestions for improvement.

Moreover, AI creates a non-judgmental learning environment, which is particularly beneficial for language learners who may feel anxious about making mistakes. This safe space encourages active participation and experimentation, gradually building confidence in learners' language skills.

The role of teachers in this AI-enhanced landscape is not diminished but instead redefined. With AI handling routine tasks such as basic assessments and progress tracking, educators focus on more complex aspects of language instruction. This includes facilitating nuanced discussions, addressing intricate language concepts, and providing crucial emotional support in the learning process.

10.5.2 Global Access to Quality Education

The absence of access to high-quality language education represents a significant challenge in numerous developing nations. The rise of AI has opened the door to substantial advancements in global access to top-notch language education. AI-driven education tools transcend geographical, socioeconomic, and linguistic barriers, promoting equitable information dissemination. However, addressing the technological hurdles of delivering AI-based education to remote areas is essential. Next-generation communication networks like 6G are pivotal in realizing this vision (Kamalov et al., 2023).

10.6 AI-Enhanced Language Learning Platforms

In language learning, AI-mediated platforms have emerged as powerful tools that simplify the comprehension of intricate subjects and offer rational and user-friendly explanations, elevating the overall learning experience. These platforms offer many opportunities, each distinguished by unique features and advantages. For instance:

Duolingo: Utilizes AI algorithms to personalize the learning journey. It provides real-time feedback on pronunciation and grammar while incorporating gamification elements to enhance engagement.
Babbel: Specializes in personalized guidance on vocabulary and grammar usage. It incorporates interactive elements, including speech recognition, enabling active practice of language skills.
Lingvist: Adapts content to meet specific learner needs, offering tailored feedback on progress. It integrates gamification strategies and supplements learning with real-world examples for practical skill development.
Busuu: Offers comprehensive, personalized feedback on vocabulary, grammar, and pronunciation. Its interactive features, including speech recognition, allow learners to refine their language abilities effectively.
Rosetta Stone: Employs advanced AI, particularly NLP, to deliver personalized feedback on pronunciation and grammar. Its interactive elements facilitate practice and enhancement of language proficiency.

Table 10.1 provides a comparative overview of these AI-mediated language learning platforms.

Table 10.1 AI-mediated language learning platforms

Platform	Application
duolingo	• Utilizes AI algorithms to personalize the learning journey • Provides real-time feedback on pronunciation and grammar • Incorporates gamification elements to enhance engagement
+Babbel	• Specializes in personalized vocabulary and grammar guidance • Incorporates interactive elements like speech recognition • Enables active practice and enhancement of language skills
LINGVIST	• Personalized content adaptation for individual learner needs • Tailored feedback on progress • Gamification strategies for engagement • Real-world examples of practical language skills
Busuu a Chegg service	• Comprehensive, personalized feedback on vocabulary, grammar, and pronunciation • Interactive speech recognition for refining language abilities
Rosetta Stone	• Advanced AI, especially NLP, for personalized pronunciation and grammar feedback • Interactive speech recognition for practicing and enhancing language proficiency

10.6.1 The Transformative Impact of AI in Education

AI significantly enhances language learning environments, improving education quality and student engagement. For example, AI-powered AR technology introduces immersive visual learning experiences, transforming traditional language textbook content into an interactive playground.

AI software plays a crucial role in personalizing language learning experiences. It analyzes students' study habits and creates personalized learning paths accordingly. This is particularly beneficial in language classrooms where teachers cannot observe each student individually. AI-powered educational platforms adapt to students' needs, identifying knowledge gaps, directing students to appropriate materials, and creating personalized language learning content.

In the context of language learning, AI technologies offer significant support for accessibility. For instance, AI-powered speech recognition and text-to-speech technologies assist language learners with hearing or visual impairments, allowing them to participate effectively in language classes and assessments.

The pandemic highlighted the importance of AI in education, particularly for remote language learning. Modern AI chatbots respond rapidly to student inquiries about language assignments, vocabulary, and grammar, enhancing the remote learning experience.

AI's role extends to language essay grading and plagiarism detection. Educational organizations use AI to assess language essays and assign grades, while services like Turnitin utilize AI to detect plagiarism in student work, ensuring academic integrity in language courses.

These AI-enhanced language learning platforms and technologies collectively represent a dynamic and evolving landscape, catering to diverse learning needs and preferences while leveraging technology to optimize the language learning process.

10.7 The Metaverse as a Language Learning Space

This section explores the metaverse as a platform for language education, focusing on its diverse aspects, unique features, and benefits.

10.7.1 Introduction to Metaverse in Education

The field of education has witnessed substantial transformations in recent years, with the COVID-19 pandemic accelerating the shift toward virtual learning environments. The metaverse has emerged as a critical area of interest for educators worldwide,

particularly in language education. Its objective is to engage students in immersive experiences that enhance the effectiveness and entertainment value of learning (Imannezhad et al., 2023; Kye et al., 2021; Li & Yu, 2023).

10.7.2 The Concept of Metaverse in Language Education

The metaverse in education can be conceptualized as an enriched educational environment that integrates metaverse-related technologies, seamlessly merging aspects from both virtual and real-world educational settings. It empowers learners to access language learning environments without constraints related to time and location. Learners utilize digital identities for immediate interactions with various entities, such as avatars, intelligent NPCs, or virtual learning resources, experiencing a strong sense of presence akin to being physically present in a real-world language classroom (Li & Yu, 2023).

10.7.3 Educational Immersion in the Metaverse Language Learning

The metaverse provides enhanced immersive experiences within virtual space and fosters interactive language learning opportunities. Students engage in face-to-face interactions within a virtual language classroom without needing physical presence. Leveraging avatars within the metaverse opens up opportunities for authentic language practice and cultural exchange within the virtual environment (Pradana & Elisa, 2023).

10.8 Features of Metaverse in Language Education

This section explores key features of the metaverse in language learning contexts, highlighting its potential to transform traditional language education paradigms.

10.8.1 Enhanced Creative Freedom and Sharing in Language Learning

The metaverse offers unprecedented creative freedom in language learning environments. It allows learners to engage in diverse real-world language scenarios, from casual conversations to professional interactions, and experience language

immersion in virtual cultural settings, transcending geographical limitations. Users create and share language learning content, fostering a global collaborative learning ecosystem. This environment supports self-directed language acquisition, encouraging critical thinking, cultural understanding, and innovative approaches to language mastery.

10.8.2 Fostering Immersive Language Learning Through Virtualization

Metaverse-based language education surpasses traditional online learning by offering enhanced interaction mimicking real-world language use. It provides unlimited virtual spaces for diverse linguistic and cultural experiences, overcoming physical constraints in language practice. The metaverse incorporates innovative technologies for personalized language learning experiences.

Platforms like Gathertown, as shown in Fig. 10.4, exemplify these features. They provide real-time multilingual communication, seamless integration of learning resources, and customizable virtual environments for targeted language acquisition. These features create interactive language learning environments that significantly enhance engagement and effectiveness compared to conventional online courses.

The "Classroom Map" feature in Zepeto, as shown in Fig. 10.5, gained significant popularity during school closures. This virtual space allowed students to meet and interact with friends in a familiar setting, mimicking the social aspects of physical classrooms, and participate in virtual lessons and activities, maintaining a sense of educational continuity.

Platforms like Roblox have also contributed to this trend. Introduced in 2006, Roblox offers a virtual currency system ("Robux") that supports a thriving economic

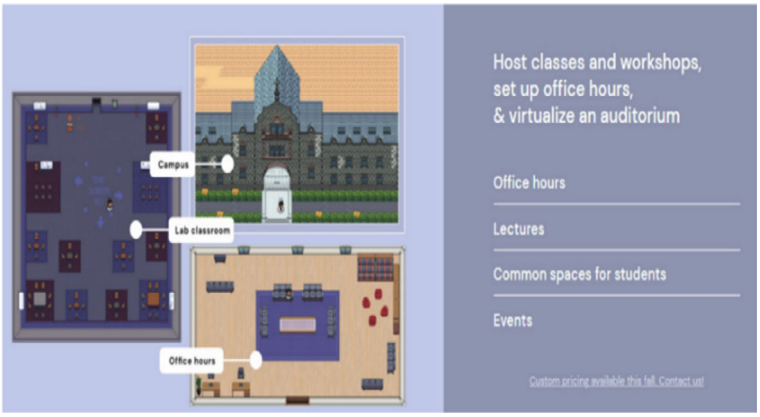


Fig. 10.4 Classroom in Gathertown (Kye et al., 2021, Fig. 6)

Fig. 10.5 Zepeto classroom (Kye et al., 2021, Fig. 7)

ecosystem, user-created spaces and real-time gaming experiences, and opportunities for users to craft VR games or participate in games designed by others. These features, illustrated in Fig. 10.6 (A-D), make Roblox not just a gaming platform but a potential educational tool, allowing for immersive learning experiences (Kye et al., 2021).

10.8.3 Metaverse Advantages for Language Learning

In the twenty-first century, language learning has become essential for developing multilingual proficiency. Traditional language learning methods often lack sufficient contextual practice and interaction, resulting in passive learning experiences. Researchers have pointed out that the advent of the metaverse holds significant potential for enhancing language learning (X. Zhang et al., 2022).

The metaverse offers several advantages for language learning:

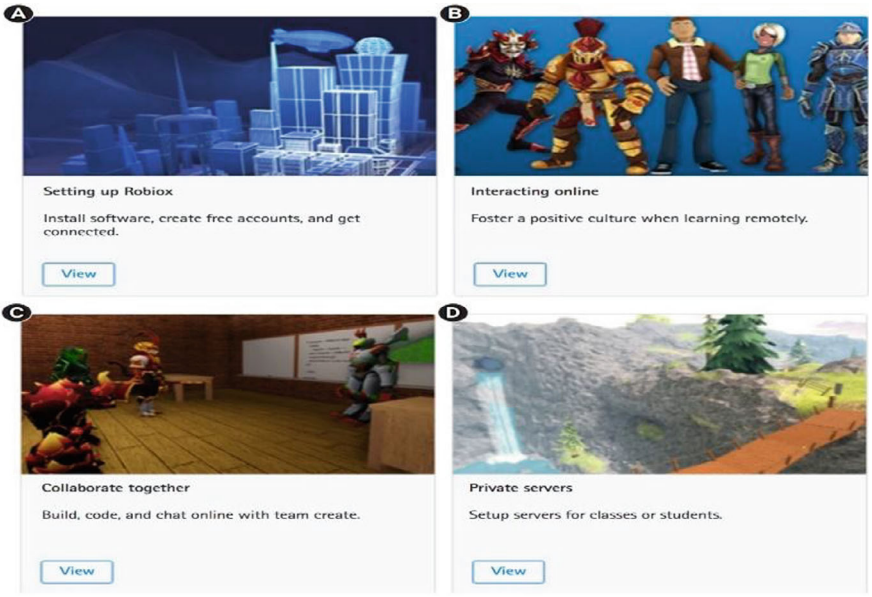


Fig. 10.6 (A-D) Roblox remote education tool

1. **Continuous Practice**: Learners invite language partners to enter virtual worlds through avatars, practicing dialogue within pre-existing or recreated environments, even when physically distant. If human partners are unavailable, intelligent language agents provide additional practice opportunities (X. Zhang et al., 2022).
2. **Immersive Experiences**: The metaverse immerses English as a Foreign Language (EFL) learners in an alternate life, providing an environment where they use the target language for work, education, socializing, and entertainment, akin to the experiences of native speakers (Chien et al., 2020).
3. **AI Integration**: The metaverse opens new roles for intelligent NPC tutors, peers, and tutees. This presents opportunities for analyzing students' behaviors and interaction patterns within the metaverse, correlating them with performance levels, and exploring novel applications in language learning (Hwang & Chien, 2022).

10.8.4 Education-Based Metaverse for the Blended Learning

The term "education-based metaverse" refers to the use of metaverse-based platforms to enhance educational activities, particularly in language learning. These platforms create immersive environments that cultivate essential skills among

Fig. 10.7 Metaverse application in diverse sectors (Bailenson, 2021; X. Zhang et al., 2022, Fig. 2)

learners, including effective collaboration, communication proficiency, and creative problem-solving abilities in the target language (Li & Yu, 2023).

The immersive technology of the metaverse introduces new possibilities for blended learning in language education. Education-focused metaverse applications like virtual classrooms seamlessly integrate with in-person learning, combining online resources with metaverse-based tools within a physical classroom (X. Zhang et al., 2022).

A notable example is the "Virtual People" course introduced by Professor Jeremy Bailenson at Stanford University in 2021. This course adopted a blended learning approach, with students equipped with VR headsets convening in the metaverse via Zoom and ENGAGE platforms. Throughout the metaverse learning sessions, participants engaged in various activities, including virtual field trips, group discussions, quizzes, and the collaborative creation of virtual spaces (Bailenson, 2021). Figure 10.7 illustrates a class discussion segment within this course conducted in the metaverse platform.

10.8.5 Examples of Metaverse Applications in Language Education

The metaverse offers a broad spectrum of potential applications in education, including language learning, medical training, and science education (Díaz et al., 2020; Hwang & Chien, 2022; Jagatheesaperumal et al., 2022; Koo, 2021; Siyaev & Jo, 2021).

Several technology companies are exploring the potential applications of metaverse technologies in language education:

1. **Roblox**: This platform offers a multiplayer game system that enhances learners' social interaction experiences. Its framework allows users significant creative

Table 10.2 Exemplary and productive cases in metaverse education at tech companies

Tech Companies	Features	Applications
ROBLOX POWERING IMAGINATION™	• User-generated content • Multiplayer interactions	• Language immersion through user-created scenarios • Peer-to-peer language practice
mondly languages	• VR language lessons • speech recognition	• Immersive vocabulary and conversation practice • Real-time pronunciation feedback
ImmerseMe	• VR cultural experiences • Interactive language scenarios	• Context-based language learning • Cultural immersion for language acquisition
XIRANG	• Long-term data storage capabilities • Multi-device support	• VR classrooms for K-12 education • VR laboratories for higher education

freedom, enabling them to enrich educational content through their own creations. In November 2021, Roblox announced a substantial initiative to invest over $10 million in non-profit organizations and to develop three educational games.

2. **Mondly**: Specializing in VR language learning, Mondly provides immersive experiences that facilitate "hands-on" language practice in simulated real-world environments.
3. **ImmerseMe**: This application enables users to practice languages in VR scenarios. Utilizing AR technology, ImmerseMe integrates digital language content with real-world environments, offering learners the opportunity to practice languages in contextually relevant settings.

Table 10.2 presents a summary of noteworthy implementations of metaverse technologies in language education by technology companies:

10.9 Industrial Perspectives on AI in Language Education

The integration of AI into language education represents a burgeoning field of research and development. These technologies demonstrate significant potential for transforming educational methodologies, with numerous initiatives currently underway to implement AI-driven educational projects.

10.9.1 AI in Education Market Size

Recent studies indicate substantial growth projections for the global education technology market. Jane Aston et al. (2022) report an anticipated compound annual growth rate (CAGR) of 17% from 2021 to 2025. This growth is attributed to an increasing demand for personalized learning experiences and enhanced accessibility of digital content. Corroborating these findings, Sharma and Mohan (2022) project that the global e-learning market will reach a valuation of $374.3 billion by 2026, driven by a CAGR of 14.6% from 2021 to 2026. The authors identify several key factors contributing to this growth, including the proliferation of mobile and digital learning platforms, a growing emphasis on skill-oriented education, and the widespread integration of AI within educational frameworks.

Emerging technologies such as AI are impacting nearly every industry, with education being no exception. The Gartner Hype Cycle for Education highlights innovations that can be leveraged to address challenges and create new opportunities to improve student learning and impact educational outcomes. This trend is reflected in Gartner's analysis of the AI software market in education, which provides further insight into AI's growing influence. Gartner predicts a 17.3% increase in global spending on AI software in the education sector by 2024, reaching $7.7 billion. Looking further ahead, they forecast a CAGR of 18.5% over the subsequent five-year period, with the potential market value reaching $13.2 billion by 2027 (Fig. 10.8).

The Gartner Hype Cycle for Education, illustrated in Fig. 10.9, shows how various AI technologies are positioned in terms of expectations and maturity, helping educators and policymakers understand which innovations might be most relevant and impactful in the near and long term.

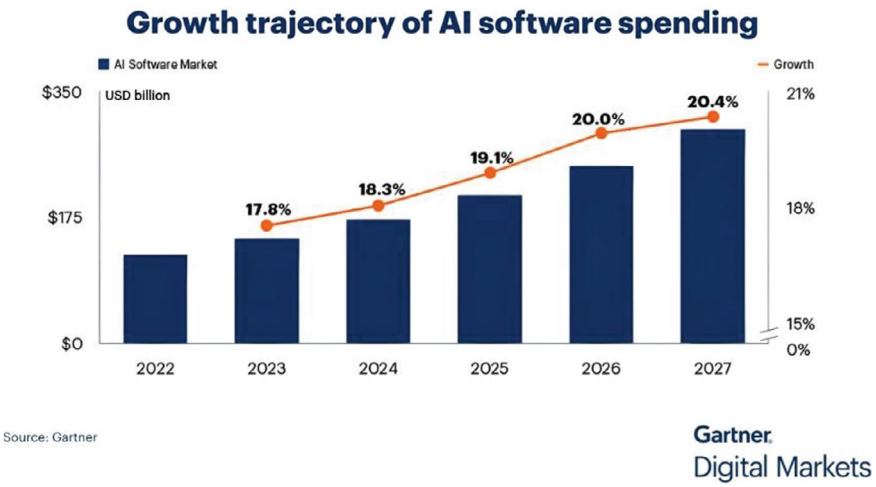


Fig. 10.8 AI software forecast ($m) and Growth (%) (Gartner, 2024a, Fig. 3)

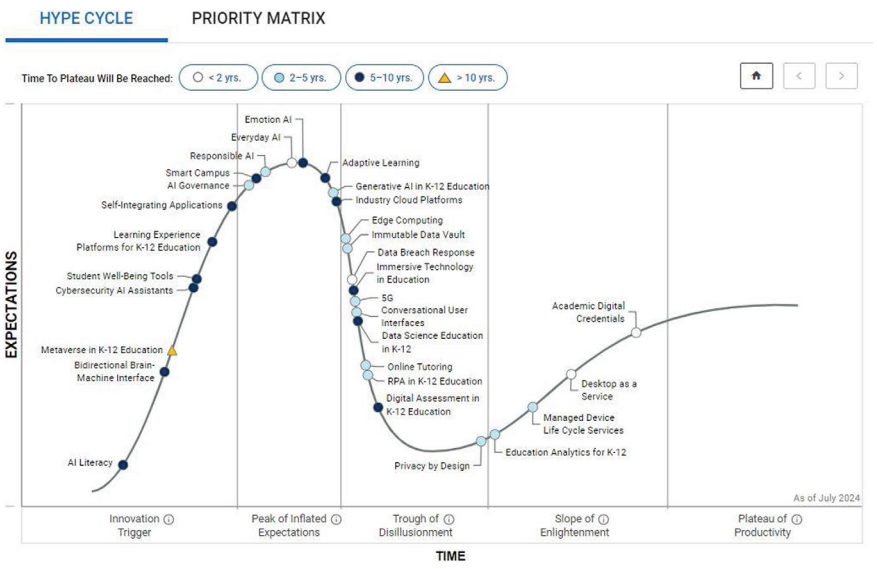


Fig. 10.9 Hype cycle for education, 2024

These projections and analyses underscore the significant role AI is expected to play in shaping the future of education, particularly in addressing challenges and enhancing learning outcomes. As these technologies mature and find practical applications, they have the potential to transform traditional educational approaches and create more personalized, efficient, and effective learning environments.

10.9.2 Hype Cycle for Artificial Intelligence, 2024: Implications for Education

While the Gartner Hype Cycle for Education provides insights specific to the education sector, examining the broader Hype Cycle for AI, 2024 offers valuable context for understanding how general AI trends may influence language education and educational technology.

The 2024 AI Hype Cycle reveals several key trends with potential implications for education. Notably, Generative AI (GenAI) has moved past the Peak of Inflated Expectations, indicating a shift toward more practical applications. In language education, this could lead to more sophisticated AI-driven language generation tools, personalized content creation, and adaptive learning systems. The recommendation to base future system architectures on composite AI techniques suggests that language education platforms may increasingly combine multiple AI approaches, resulting in more comprehensive and effective language learning tools.

Growing attention to AI governance, risk management, and safety aligns with the educational sector's need for responsible AI implementation. This trend may influence the development of AI-powered educational tools that prioritize ethical considerations and student data protection. The advancement of AI engineering and knowledge graphs could enhance the scalability and robustness of AI solutions in education, potentially improving the representation and connection of language concepts in learning systems.

Technologies at the Innovation Trigger stage, such as AI-ready data, causal AI, and multiagent systems, may pave the way for next-generation educational AI, enabling more sophisticated language understanding, improved educational decision-making, and more interactive learning environments. Despite some challenges, the progress in cloud AI services could make powerful AI tools more accessible to educational institutions, potentially democratizing access to advanced language learning technologies.

These trends suggest that the language education sector may see more sophisticated, ethical, and accessible AI-powered tools in the coming years. Educators and institutions should prepare for a landscape where AI becomes increasingly integral to language teaching and learning, while also remaining mindful of the need for responsible implementation and the potential for rapid technological change. The Gartner Hype Cycle for AI is illustrated in Fig. 10.10.

A comprehensive study, conducted by Technavio (2023), offers additional perspectives on the AI market within the education sector. Their findings suggest

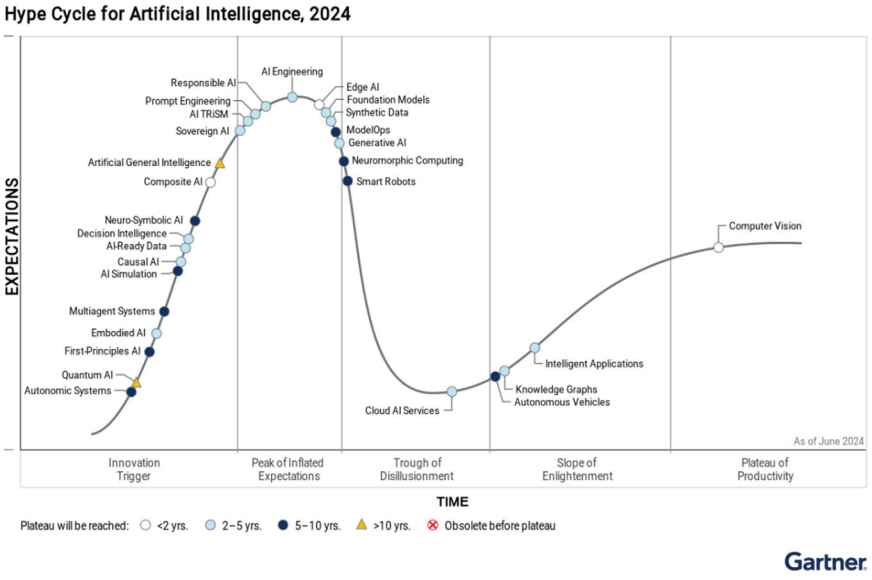


Fig. 10.10 Hype Cycle for Artificial Intelligence, 2024 (Gartner, 2024b, Fig. 1)

Fig. 10.11 Global AI market 2023–2027 in the education sector (Technavio, 2023, Fig. 1)

a CAGR of 41.14% from 2022 to 2027, with an anticipated market expansion of USD 1,100.07 million (Fig. 10.11).

It is noteworthy that these substantial financial investments are complemented by high-profile policy statements from international organizations such as the European Commission (EC), the Organization for Economic Co-operation and Development (OECD), and the United Nations Educational, Scientific and Cultural Organization (UNESCO). These statements underscore the importance of both promoting and regulating these emerging technologies (Holmes & Tuomi, 2022; Mou, 2019).

The involvement of these major international bodies highlights the global recognition of AI's potential impact on education. It also emphasizes the need for careful consideration of ethical, legal, and social implications as AI becomes more prevalent in educational settings.

10.10 Potential Challenges and Future Directions

While AI and the metaverse offer innovative prospects for language learning, it's essential to remain vigilant regarding various challenges in their implementation. This section will delve into the key challenges posed by AI in AIEd and the metaverse, specifically in the context of language learning.

10.10.1 Bias and Discrimination

The utilization of AIEd holds great promise for improving teaching and learning experiences. However, bias and discrimination pose a substantial challenge to its fair

and effective implementation. Bias in AI systems primarily arises from the training data, which often carries linguistic and cultural biases. Developers and users of these algorithms also inadvertently reinforce discrimination against marginalized groups. Generative AI further perpetuates cultural stereotypes, biases, and prejudiced language, causing harm to marginalized communities such as racial minorities, women, students with disabilities, and LGBTQ+ individuals.

The impact of these biases on education is significant, potentially leading to biased classroom outcomes and perpetuating or worsening existing inequalities. This affects students' learning outcomes, opportunities, and access to resources. For instance, AI tools often struggle to recognize the dialects and accents of non-native English-speaking students, leading to misunderstandings. Facial recognition AI frequently fails to accurately identify students with darker skin tones, resulting in skewed data and potential bias. Additionally, AI-generated content exhibits cultural biases, further exacerbating these issues in educational settings.

Since most of the internet content originates from Global North, AI models trained on this data are shaped by those perspectives. To address the AI divide identified in the Oxford Insights Government AI Readiness Index, it is essential to further tune and customize AI for non-Western educational institutions to align with regional and local cultural norms. In the context of assessment and evaluation, AI tools may demonstrate bias toward specific linguistic or cultural backgrounds, potentially impacting the assessment outcomes for students from diverse backgrounds (Kamalov et al., 2023).

To address these challenges, several strategies are employed. These include ensuring diverse and representative training data, implementing algorithmic transparency and explainability, maintaining human oversight in AI decision-making processes, developing and adhering to ethical AI guidelines and frameworks, and regularly auditing AI systems for bias and discrimination. Addressing these challenges necessitates increased collaboration among software engineers, educators, and policymakers. Together, they work to create transparent, accountable, and inclusive AI systems that ensure the responsible and equitable deployment of AI in language education.

10.10.2 Ethical Considerations

The integration of AI in education brings new ethical implications and risks. For instance, administrators may be tempted to replace human teaching with profitable automated AI solutions during budget cuts. Such a transition could lead to concerns among faculty members, teaching assistants, student counselors, and administrative staff, who may fear that intelligent tutors, expert systems, and chatbots could replace their roles and jobs (Zawacki-Richter et al., 2019). The critical and urgent need for AI ethics in education calls for collaborative efforts from all stakeholders, including educators, administrators, researchers, technology innovators, and members of society at large (K. Zhang & Aslan, 2021).

Several institutions have recently emerged, intending to establish an ethical governance framework for AIEd. These include the Institute for Ethical AIEd in the UK, the European Commission Ethics guidelines for trustworthy AI, and the upcoming UNESCO IESALC Manual on AI in higher education. Similarly, the Analysis & Policy Observatory released a discussion paper in April 2019, intending to establish an AI ethics framework for Australia.

These efforts are crucial in navigating the complexities of AIEd and ensuring its integration benefits everyone involved.

10.10.3 Data Privacy Security

Privacy and security concerns are gaining prominence as personal and sensitive information is increasingly shared and stored online. This poses a significant challenge for educational institutions, such as schools and colleges, which must prioritize the safety and security of student data. The growing threat of cyberattacks and data breaches underscores the need for these institutions to implement robust measures to safeguard student information (Ahmad et al., 2023).

Multiple solutions have been suggested to address these concerns, including the adoption of comprehensive data protection policies, the utilization of encryption methods, and the enforcement of strict access controls. It is imperative for educational institutions to prioritize the safeguarding of sensitive data and adhere to data privacy regulations like the General Data Protection Regulation (GDPR) to uphold trust and mitigate the risk of potential misuse. Furthermore, promoting a culture of privacy awareness within educational institutions is essential to guarantee adherence to data protection laws and regulations. It necessitates the establishment of comprehensive guidelines for educators and AI developers, delineating best practices and emphasizing transparency, accountability, and ethical AI usage in educational settings. Eventually, addressing data privacy and security concerns in AI-enhanced education is paramount to maximizing the advantages of these technologies while mitigating potential risks (Kamalov et al., 2023).

10.10.4 Metaverse Challenges in Education

Integrating the metaverse into language education presents unique challenges. It demands substantial investments in technical infrastructure, including high-speed internet access and VR/AR headsets, which may not be available to all language learners globally. Developing immersive language learning experiences in the metaverse is challenging and time-intensive. Not all students may have access to the required technology, potentially exacerbating existing inequalities in language education.

While the metaverse offers opportunities for immersive cultural exchanges, managing these interactions ethically and safely poses challenges. As the number of language learners in metaverse environments grows, technical issues like latency and server strain may arise (Ali, 2020). The metaverse, being a dynamic and swiftly advancing technology, currently lacks standardized practices and seamless compatibility across various platforms and devices, creating challenges in developing consistent language learning experiences.

Addressing these challenges requires collaboration between language education experts, technologists, and policymakers to create ethical, accessible, and effective AI and metaverse-based language learning environments.

10.11 Future Research Directions

Recent studies have demonstrated that language learners generally hold positive attitudes toward AI tools designed for language learning (H. H.-J. Chen et al., 2023; Moussalli & Cardoso, 2020; Wang et al., 2022). AI technologies offer immediate feedback and enhance the learning environment flexibility, enabling learners to gain greater independence and access more learning opportunities beyond the classroom.

10.11.1 Critical Gaps in AI and Metaverse Language Learning Research

A comprehensive review of studies spanning from 2015 to 2021 examining the incorporation of AI in language learning and teaching has revealed several critical gaps (Sharadgah & Sa'di, 2022). These gaps encompass concerns related to body language, gestures, expressions, emotions, translation, the absence of detailed descriptions of teaching materials used in AI-driven learning, and uncertainties regarding the scope of AI applications. In the context of metaverse-based language learning, additional gaps include the lack of research on long-term effectiveness, cultural adaptation in virtual environments, and the impact on language acquisition compared to traditional methods.

10.11.2 Exploring Promising AI and Metaverse Research Areas

Future research efforts should focus on the convergence of AI and metaverse technologies in language education. This includes investigating AI-powered virtual language tutors within metaverse environments, exploring the effectiveness of

immersive cultural exchanges for language acquisition, and developing AI systems that adapt to learners' progress in real time within virtual worlds. Collaborative interdisciplinary research involving educators, technologists, and specialists in emerging technologies is crucial for comprehensive insights. The scope of research should expand to examine AI and metaverse applications in language education across institutional, regional, and national levels over extended periods.

10.11.3 Exploring Emerging Technologies in AI and Metaverse Language Learning

It is crucial to explore the potential of emerging technologies that further enhance language learning in the metaverse. VR, AR, and MR have the potential to revolutionize language education by providing immersive, context-rich environments for language practice. Research should focus on how these technologies can be integrated with AI to create adaptive, personalized learning experiences within metaverse platforms.

10.11.4 Empowering Educators for Effective AI and Metaverse Integration

There is growing recognition that educators should be equipped with the necessary skills to leverage AI and metaverse technologies effectively. Future research should prioritize teacher training and empowerment, focusing on developing educators' fundamental knowledge of AI and metaverse technologies, as well as proficiency in their practical application to enhance teaching effectiveness and address students' educational needs in virtual environments.

10.11.5 Ethical Considerations and Teacher Preparedness

Future research and practices must actively engage with the potential and challenges associated with AI and metaverse pedagogical and technical development, ensuring their effective and ethical utilization. This includes addressing concerns about language teachers' preparedness for AI and metaverse integration, data privacy in virtual learning environments, and ethical considerations when conducting research involving AI, metaverse technologies, and learner/educator data.

By addressing these research directions, we work toward a more comprehensive understanding of how AI and the metaverse revolutionize language learning, while also ensuring ethical and effective implementation.

10.12 Conclusion

This chapter underscores the transformative impact of diverse AI technologies in language education, particularly in the context of the metaverse era. These advancements are not just theoretical constructs; they are actively reshaping how we approach language learning, offering more interactive and immersive educational experiences.

We have endeavored to bridge the gap between cutting-edge AI innovations and their real-world applications in educational settings. This involved gathering insights from a broad range of academic and industry sources, highlighting current trends and potential uses of AI in language education.

As we look to the future, the integration of AI into language learning within the metaverse holds great promise for making education more inclusive and accessible. This shift toward a more equitable educational landscape presents challenges, but the potential benefits are significant.

In conclusion, we recognize that our research may not capture every aspect of this rapidly evolving field. We anticipate that our findings will inspire further research, particularly in maximizing the potential of AI to foster equitable language education in the metaverse.

References

Abilkaiyrkyzy, A., Elhagry, A., Laamarti, F., & Elsaddik, A. (2023). Metaverse key requirements and platforms survey. *IEEE Access, 11*, 117765–117787.

Ahmad, S., Umirzakova, S., Mujtaba, G., Amin, M. S., & Whangbo, T. (2023). *Education 5.0: requirements, enabling technologies, and future directions.* arXiv preprint arXiv:2307.15846

Aldosari, S. A. M. (2020). The future of higher education in the light of artificial intelligence transformations. *International Journal of Higher Education, 9*(3), 145–151.

Alfalah, A. A. (2023). Factors influencing students' adoption and use of mobile learning management systems (m-LMSs): A quantitative study of Saudi Arabia. *International Journal of Information Management Data Insights, 3*(1), Article 100143.

Ali, Z. (2020). *Artificial intelligence (AI): A review of its uses in language teaching and learning.* Paper presented at the IOP Conference Series: Materials Science and Engineering (Vol. 769). IOP Publishing.

Anderson, C. R. (2023). *Traditional versus modern teaching methods amongst special education students and enhancing students' self-concept: A comprehensive literature review.* The Chicago School of Professional Psychology.

Andriulli, F., Chen, P.-Y., Erricolo, D., & Jin, J.-M. (2022). Guest editorial machine learning in antenna design, modeling, and measurements. *IEEE Transactions on Antennas and Propagation, 70*(7), 4948–4952.

Bailenson, J. (2021). *COMM 166/266: Virtual people.* https://vhil.stanford.edu/downloads/comm166

Baker, T., & Smith, L. (2019). *Educ-AI-tion rebooted? Exploring the future of artificial intelligence in schools and colleges.* https://media.nesta.org.uk/documents/Future_of_AI_and_education_v5_WEB.pdf

Baskara, R. (2023). Exploring the implications of ChatGPT for language learning in higher education. *Indonesian Journal of English Language Teaching and Applied Linguistics, 7*(2), 343–358.

Bates, T., Cobo, C., Mariño, O., & Wheeler, S. (2020). *Can artificial intelligence transform higher education?* (Vol. 17, pp. 1–12). SpringerOpen.

Bowles, D. C., & Kruger, J. S. (2023). *Generating employable, intelligent graduates in a world with Generative AI: Thoughts for educators* (Vol. 9, pp. 75–77). Sage.

Carbonell, J. R. (1970). AI in CAI: An artificial-intelligence approach to computer-assisted instruction. *IEEE Transactions on Man-Machine Systems, 11*(4), 190–202.

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

Chen, X., Xie, H., Zou, D., & Hwang, G.-J. (2020). Application and theory gaps during the rise of Artificial Intelligence in Education. *Computers and Education: Artificial Intelligence, 1*, Article 100002.

Chien, S.-Y., Hwang, G.-J., & Jong, M.S.-Y. (2020). Effects of peer assessment within the context of spherical video-based virtual reality on EFL students' English-Speaking performance and learning perceptions. *Computers & Education, 146*, Article 103751.

Cline, E. (2011). *Ready player one.* Ballantine Books.

Collins, C. (2008). Looking to the future: Higher education in the Metaverse. *Educause Review, 43*(5), 50–52.

Díaz, J., Saldaña, C., & Avila, C. (2020). Virtual world as a resource for hybrid education. *International Journal of Emerging Technologies in Learning (iJET), 15*(15), 94–109.

Divekar*, R. R., Drozdal*, J., Chabot*, S., Zhou, Y., Su, H., Chen, Y., Zhu, H., Hendler, J. A., & Braasch, J. (2022). Foreign language acquisition via artificial intelligence and extended reality: Design and evaluation. *Computer Assisted Language Learning, 35*(9), 2332–2360. https://doi.org/10.1080/09588221.2021.1879162

Edwards, B. I., & Cheok, A. D. (2018). Why not robot teachers: Artificial intelligence for addressing teacher shortage. *Applied Artificial Intelligence, 32*(4), 345–360.

Ekin, C. C., Polat, E., & Hopcan, S. (2023). Drawing the big picture of games in education: A topic modeling-based review of past 55 years. *Computers & Education, 194*, Article 104700.

Elme, L., Jørgensen, M. L., Dandanell, G., Mottelson, A., & Makransky, G. (2022). Immersive virtual reality in STEM: Is IVR an effective learning medium and does adding self-explanation after a lesson improve learning outcomes? *Educational Technology Research and Development, 70*(5), 1601–1626.

Fullan, M. (2023). *The principal 2.0: Three keys to maximizing impact*. Wiley.

Gartner. (2024a). *Compare AI software spending in the education industry, 2023–2027*. https://www.gartner.com/en/documents/5238163

Gartner. (2024b). *Hype cycle for artificial intelligence, 2024*. https://www.gartner.com/en/documents/5505695

Godwin-Jones, R. (2017). Smartphones and language learning. *Language Learning & Technology, 21*(2), 3–17.

Grassini, S. (2023). Shaping the future of education: Exploring the potential and consequences of AI and ChatGPT in educational settings. *Education Sciences, 13*(7), 692. https://www.mdpi.com/2227-7102/13/7/692

Guan, C., Mou, J., & Jiang, Z. (2020). Artificial intelligence innovation in education: A twenty-year data-driven historical analysis. *International Journal of Innovation Studies, 4*(4), 134–147.

Hinojo-Lucena, F.-J., Aznar-Díaz, I., Cáceres-Reche, M.-P., & Romero-Rodríguez, J.-M. (2019). Artificial Intelligence in higher education: A bibliometric study on its impact in the scientific literature. *Education Sciences, 9*(1), 51. https://www.mdpi.com/2227-7102/9/1/51

Holmes, W., & Tuomi, I. (2022). State of the art and practice in AI in education. *European Journal of Education, 57*(4), 542–570.

Hong, W. C. H. (2023). The impact of ChatGPT on foreign language teaching and learning: Opportunities in education and research. *Journal of Educational Technology and Innovation, 5*(1), 37–45.

Hu, J., & Hu, X. (2020). *The effectiveness of autonomous listening study and pedagogical implications in the module of artificial intelligence.* Paper presented at the Journal of Physics: Conference Series.

Huang, T.-H., & Wang, L.-Z. (2023). Artificial intelligence learning approach through total physical response embodiment teaching on French vocabulary learning retention. *Computer Assisted Language Learning, 36*(8), 1608–1632.

Huang, X., Zou, D., Cheng, G., Chen, X., & Xie, H. (2023). Trends, research issues and applications of artificial intelligence in language education. *Educational Technology & Society, 26*(1), 112–131.

Huh, S. (2023). Are ChatGPT's knowledge and interpretation ability comparable to those of medical students in Korea for taking a parasitology examination? A descriptive study. *Journal of Educational Evaluation for Health Professions, 20*(1), 1.

Hwang, G.-J., & Chien, S.-Y. (2022). Definition, roles, and potential research issues of the metaverse in education: An artificial intelligence perspective. *Computers and Education: Artificial Intelligence, 3*, Article 100082.

Ijaz, K., Bogdanovych, A., & Trescak, T. (2017). Virtual worlds vs books and videos in history education. *Interactive Learning Environments, 25*(7), 904–929.

Imannezhad, S., Vahedian-Shahroodi, M., Shariati, K., Mansourzadeh, A., & Saeedi, M. (2023). Metaverse in education; an overview of systematic reviews. *Medical Education Bulletin, 4*(2), 731–743. https://doi.org/10.22034/meb.2023.405301.1079

Jagatheesaperumal, S. K., Ahmad, K., Al-Fuqaha, A., & Qadir, J. (2022). Advancing education through extended reality and internet of everything enabled metaverses: Applications, challenges, and open issues. *arXiv preprint*arXiv:2207.01512

Jane Aston, E. D., Maria Guijon, K. L., & Popov, D. (2022). *The education technology market in England*. https://assets.publishing.service.gov.uk/media/636e7717e90e07186280f7cf/Edtech_market_in_England_Nov_2022.pdf

Kabudi, T., Pappas, I., & Olsen, D. H. (2021). AI-enabled adaptive learning systems: A systematic mapping of the literature. *Computers and Education: Artificial Intelligence, 2*, Article 100017.

Kamalov, F., Santandreu Calonge, D., & Gurrib, I. (2023). New era of artificial intelligence in education: Towards a sustainable multifaceted revolution. *Sustainability, 15*(16), 12451. https://www.mdpi.com/2071-1050/15/16/12451

Kemp, J., & Livingstone, D. (2006). *Putting a second life "metaverse" skin on learning management systems.* Paper presented at the Proceedings of the Second Life education workshop at the Second Life community convention.

Keshav, N. U., Salisbury, J. P., Vahabzadeh, A., & Sahin, N. T. (2017). Social communication coaching smartglasses: Well tolerated in a diverse sample of children and adults with autism. *JMIR mHealth and uHealth, 5*(9), Article e8534.

Khang, A., Muthmainnah, M., Seraj, P. M. I., Al Yakin, A., & Obaid, A. J. (2023). AI-Aided teaching model in education 5.0. In *Handbook of research on AI-based technologies and applications in the era of the metaverse* (pp. 83–104). IGI Global.

Kim, H., Hwang, S., Kim, J., & Lee, Z. (2022). Toward smart communication components: Recent advances in human and AI speaker interaction. *Electronics, 11*(10), 1533. https://www.mdpi.com/2079-9292/11/10/1533

Kohnke, L., Moorhouse, B. L., & Zou, D. (2023). ChatGPT for language teaching and learning. *RELC Journal, 54*(2), 537–550.

Koo, H. W. (2021). Training in lung cancer surgery through the metaverse, including extended reality, in the smart operating room of Seoul National University Bundang Hospital, Korea. *Journal of Educational Evaluation for Health Professions, 18.*

Kuhail, M. A., Alturki, N., Alramlawi, S., & Alhejori, K. (2023). Interacting with educational chatbots: A systematic review. *Education and Information Technologies, 28*(1), 973–1018. https://doi.org/10.1007/s10639-022-11177-3

Kye, B., Han, N., Kim, E., Park, Y., & Jo, S. (2021). Educational applications of metaverse: Possibilities and limitations. *Journal of Educational Evaluation for Health Professions, 18.*

Li, M., & Yu, Z. (2023). A systematic review on the metaverse-based blended English learning. *Frontiers in Psychology, 13*, 1087508.
Lim, W. M., Gunasekara, A., Pallant, J. L., Pallant, J. I., & Pechenkina, E. (2023). Generative AI and the future of education: Ragnarök or reformation? A paradoxical perspective from management educators. *The International Journal of Management Education, 21*(2), Article 100790.
Liu, M. (2023). *Exploring the application of artificial intelligence in foreign language teaching: Challenges and future development.* Paper presented at the SHS Web of Conferences.
Liu, Y., Gu, J., Goyal, N., Li, X., Edunov, S., Ghazvininejad, M., Lewis, M., & Zettlemoyer, L. (2020). Multilingual denoising pre-training for neural machine translation. *Transactions of the Association for Computational Linguistics, 8*, 726–742.
Lv, Z. (2023). Generative artificial intelligence in the metaverse era. *Cognitive Robotics, 3*, 208–217.
Mageira, K., Pittou, D., Papasalouros, A., Kotis, K., Zangogianni, P., & Daradoumis, A. (2022). Educational AI Chatbots for content and language integrated learning. *Applied Sciences, 12*(7), 3239. https://www.mdpi.com/2076-3417/12/7/3239
Mou, X. (2019). Artificial intelligence: Investment trends and selected industry uses. *International Finance Corporation, 8*(2), 311–320.
Moussalli, S., & Cardoso, W. (2020). Intelligent personal assistants: Can they understand and be understood by accented L2 learners? *Computer Assisted Language Learning, 33*(8), 865–890.
Natarajan, B., Rajalakshmi, E., Elakkiya, R., Kotecha, K., Abraham, A., Gabralla, L. A., & Subramaniyaswamy, V. (2022). Development of an end-to-end deep learning framework for sign language recognition, translation, and video generation. *IEEE Access, 10*, 104358–104374.
Nilsson, N. J. (1998). *Artificial intelligence: A new synthesis*. Morgan Kaufmann.
Okonkwo, C. W., & Ade-Ibijola, A. (2021). Chatbots applications in education: A systematic review. *Computers and Education: Artificial Intelligence, 2*, Article 100033.
Pradana, M., & Elisa, H. P. (2023). Metaverse in education: A systematic literature review. *Cogent Social Sciences, 9*(2), 2252656. https://doi.org/10.1080/23311886.2023.2252656
Qadir, J. (2023). *Engineering education in the era of ChatGPT: Promise and pitfalls of generative AI for education.* Paper presented at the 2023 IEEE Global Engineering Education Conference (EDUCON).
Razumovskaia, E., Glavas, G., Majewska, O., Ponti, E. M., Korhonen, A., & Vulic, I. (2022). Crossing the conversational chasm: A primer on natural language processing for multilingual task-oriented dialogue systems. *Journal of Artificial Intelligence Research, 74*, 1351–1402.
Schlemmer, E. (2014). *Learning in metaverses: Co-existing in real virtuality: Co-Existing in real virtuality*. IGI Global.
Shannon, S., & Chapelle, C. (2017). *The handbook of technology and second language teaching and learning*. Wiley-Blackwell.
Sharadgah, T. A., & Sa'di, R. A. (2022). A systematic review of research on the use of artificial intelligence in English language teaching and learning (2015–2021): What are the current effects? *Journal of Information Technology Education: Research, 21*.
Sharma, R., & Mohan, M. (2022). Artificial intelligence and e-learning: Best practices across the globe. In *Adoption and implementation of AI in customer relationship management* (pp. 1–25). IGI Global.
Shortliffe, E. H., Davis, R., Axline, S. G., Buchanan, B. G., Green, C. C., & Cohen, S. N. (1975). Computer-based consultations in clinical therapeutics: Explanation and rule acquisition capabilities of the MYCIN system. *Computers and Biomedical Research, 8*(4), 303–320.
Siyaev, A., & Jo, G. S. (2021). Neuro-symbolic speech understanding in aircraft maintenance metaverse. *IEEE Access, 9*, 154484–154499. https://doi.org/10.1109/ACCESS.2021.3128616
Sleeman, D., & Brown, J. S. (1982). *Intelligent tutoring systems.* Academic Press.
Stephenson, N. (1994). *Snow crash.* Penguin UK.
Su, Y., Lin, Y., & Lai, C. (2023). Collaborating with ChatGPT in argumentative writing classrooms. *Assessing Writing, 57*, Article 100752.

Tang, K.-Y., Chang, C.-Y., & Hwang, G.-J. (2023). Trends in artificial intelligence-supported e-learning: A systematic review and co-citation network analysis (1998–2019). *Interactive Learning Environments, 31*(4), 2134–2152.

Technavio. (2023). *Artificial Intelligence (AI) market in the education sector by end-user, learning method and geography-forecast and analysis 2023–2027*. https://www.prnewswire.com/news-releases/artificial-intelligence-market-in-the-education-sector-to-grow-by-usd-1-100-07-million-from-2022-to-2027--the-increasing-emphasis-on-customized-learning-paths-using-ai-drives-market-growth---technavio-301963204.html

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, 37*(4), 814–840.

Wei, L. (2023). Artificial intelligence in language instruction: Impact on English learning achievement, L2 motivation, and self-regulated learning. *Frontiers in Psychology, 14*, 1261955.

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.

Wiederhold, B. K. (2022). *Ready (or not) player one: Initial musings on the metaverse* (Vol. 25, pp. 1–2). Mary Ann Liebert.

Wollny, S., Schneider, J., Di Mitri, D., Weidlich, J., Rittberger, M., & Drachsler, H. (2021). Are we there yet?-a systematic literature review on chatbots in education. *Frontiers in Artificial Intelligence, 4*, Article 654924.

Wright, A. M., Munz, S. M., & McKenna-Buchanan, T. (2023). Triage teaching: Exploring teacher self-efficacy during COVID-19. In *Pandemic pedagogies* (pp. 96–111). Routledge.

Xia, Q., Chiu, T. K. F., & Chai, C. S. (2023). The moderating effects of gender and need satisfaction on self-regulated learning through Artificial Intelligence (AI). *Education and Information Technologies, 28*(7), 8691–8713. https://doi.org/10.1007/s10639-022-11547-x

Yan, D. (2023). Impact of ChatGPT on learners in a L2 writing practicum: An exploratory investigation. *Education and Information Technologies, 28*(11), 13943–13967.

Yunjiu, L., Wei, W., & Zheng, Y. (2022). Artificial intelligence-generated and human expert-designed vocabulary tests: A comparative study. *SAGE Open, 12*(1), https://doi.org/10.1177/21582440221082130

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education–where are the educators? *International Journal of Educational Technology in Higher Education, 16*(1), 1–27.

Zhang, K., & Aslan, A. B. (2021). AI technologies for education: Recent research & future directions. *Computers and Education: Artificial Intelligence, 2*, Article 100025. https://doi.org/10.1016/j.caeai.2021.100025

Zhang, R., & Zou, D. (2022). Types, purposes, and effectiveness of state-of-the-art technologies for second and foreign language learning. *Computer Assisted Language Learning, 35*(4), 696–742. https://doi.org/10.1080/09588221.2020.1744666

Zhang, X., Chen, Y., Hu, L., & Wang, Y. (2022). The metaverse in education: Definition, framework, features, potential applications, challenges, and future research topics. *Frontiers in Psychology, 13*, 6063.

Muhammad Saqlain Aslam received the B.S. degree in Software Engineering from the Government College University, Faisalabad, Pakistan, in 2013, the M.S. degree in Software Engineering from the Beijing Institute of Technology, Beijing, China, in 2017, and the Ph.D. degree in Computer Science and Information Engineering from the National Central University, Taiwan, in 2021. He is currently working as an Assistant Researcher at the AI Research Center, Hon Hai (Foxconn) Research Institute, Taipei, Taiwan. His current research interests include artificial intelligence, machine learning, deep learning, biometric recognition, and computer vision.

Yung-Hui Li received the B.S. degree from the Department of Electrical Engineering, National Taiwan University, Taipei, Taiwan, in 1995, the M.S. degree from the Department of Computer and Information Science, University of Pennsylvania, Philadelphia, PA, USA, in 1998, and the Ph.D. degree from the Language Technology Institute, School of Computer Science, Carnegie Mellon University, Pittsburgh, PA, in 2010. He is currently an Associate Professor with the Department of Computer Science and Information Engineering, at National Central University, Taoyuan, Taiwan, and the Director of AI Research Center at Hon Hai Research Institute in Taipei, Taiwan. His current research interests include long-range iris recognition, including extended DoF for iris image acquisition, blur iris image segmentation, iris mask generation, and iris image super-resolution; face recognition; eye-tracking; human–computer interaction; computer vision; pattern recognition; and machine learning. Prof. Li is currently a member of the IEEE Computer Society and Signal Processing Society.

Chapter 11
AI and VR Converge: The Future of Language Learning in an Emerging Metaverse

Robert Godwin-Jones

Abstract The idea of a metaverse holds considerable promise for language learning. The phenomenon of an integrated digital and physical world, accessed in a 3D immersive environment, offers powerful opportunities for incidental language learning through informal chats, simulations, and games, as well as for dedicated language acquisition through personalized, adaptive tutoring. For that vision to be fulfilled, multimodal AI will be needed, moving beyond text-only language models, enabling input and output in any combination of media. Integrating multimodal AI into virtual reality (VR) will allow immersive experiences to be wide-ranging and free form, supplanting scripted interactions that offer limited learner agency. At the same time, memory capabilities built into AI systems will enable creation of learner profiles used to individualize interactions according to learner goals/interests and proficiency level. The copresence of peer learners, as well as the growing ability of AI systems to mimic human communicative practices, will enable a version of the metaverse, with its co-habitation of humans and AI agents, to become a collaborative space for social learning. While such a system will provide an engaging learning space, concerns of privacy and ethics will need to be addressed. For learners, it will be important to supplement AI-based virtual interactions with authentic human-to-human communication, for example, through engaging in virtual exchange. To understand the complex, intertwined interactions between humans and AI, it will be helpful to look beyond tool functionality and consider AI (and VR) from a holistic, ecological perspective, using frameworks such as sociomaterialism and complexity theory, along with insights from non-Western ontologies.

Keywords Generative AI · Metaverse · Virtual reality · Multimedia · Ecological learning theories

R. Godwin-Jones (✉)
Virginia Commonwealth University, Richmond, VA, USA
e-mail: rgjones@vcu.edu

Y. Lan et al. (eds.), *AI-Mediated Language Education in the Metaverse Era*, Chinese Language Learning Sciences, https://doi.org/10.1007/978-981-95-0245-5_11

11.1 Introduction

The notion of a metaverse, originating in Neal Stephenson's *Snow Crash* (1992), has attracted popular interest since Facebook rebranded itself as "Meta" in 2021. The concept of a merged real and virtual world, accessible principally through virtual reality (VR), seems to hold promise in many areas of human activity, including language learning. Virtual reality is poised to move to a new, higher level of functionality through the emerging integration of generative AI. For language learning, that promises to move VR apps from scripted to free-form interactions, holding the promise of highly personalized, adaptive language learning. AI is poised to become a vital partner in enabling something akin to the metaverse, particularly as its multimodal capabilities are enhanced. In the process, humans are likely to rely more heavily on AI in a variety of ways, making it into an authoring and learning partner (Godwin-Jones, 2024b).

To understand the extent of AI copresence, we will need to look beyond AI as a technological tool and consider its transforming effect in society. In that process, it will be helpful to look at ecological frameworks such as sociomaterialism and complexity theory, which analyze the intertwined dynamic among human, non-humans (including AI), and the context of use. In addition, inclusive and relational ontologies associated with Indigenous cultures provide a perspective that rejects dualism and reductionism, inviting acceptance of uncertainty and ambiguity (Canagarajah, 2021), dispositions helpful in coping with the disruption AI represents for human society.

AI-enhanced VR is likely in the long run to bring profound changes to language learning, providing exciting opportunities for learners and leading teachers to rethink traditional approaches to language learning and assessment, and possibly question the fundamental mission and goal of instructed language learning. The personalized, adaptive learning possible and practical through AI integrated into mixed reality promises to redefine the field of computer-assisted language learning (CALL). At the same time, we need to be cognizant of the many ethical issues surrounding AI creation and use. An understanding of how large language models (LLMs) are created (the basis for generative AI) and how different a phenomenon they represent from conventional software programs is important for developing a "calibrated trust" in such systems (Ranalli, 2021, p. 14), a necessary component of critical AI literacy. AI will figure largely in our lives in multiple ways, so becoming informed AI consumers will be important for us as citizens, as well as for our roles as teachers and learners.

11.2 The Changing Face of AI: Integrating Multimodality

AI is hardly a new phenomenon. A basic knowledge of its historical development can help appreciate the enormity of the success of generative AI today. Computer scientists have been working since the 1950s on ways to have machines perform

functions heretofore unique to human beings, particularly the ability to use natural language and to solve problems through something resembling human reasoning power. Since then, enthusiasm for AI has waxed and waned with research and AI going through a series of advances followed by "winters" of inactivity and lack of funding. Major breakthroughs were achieved in the early 2010s through advances in machine learning. A major innovation came in 2017 with the "transformer" architecture (Vaswani et al., 2017), which greatly enhanced performance (through parallel processing) while enabling faster and more efficient training.

A leap forward occurred in the early 2020s when the transformer architecture was applied to immense sets of data to build more effective large language models. OpenAI developed the GPT series (Generative Pre-trained Transformer) whose version 3 implementation was able to generate text in a great variety of genres and in language that closely resembles human writing (Godwin-Jones, 2021). Those text output capabilities were enhanced when the new version, 3.5, was publicly released as *ChatGPT* in late 2022. That was soon followed by other generative AI systems such as Google *Bard* (now *Gemini*), as well as by systems capable of generating images, *Midjourney* and *Stable Diffusion*. The media frenzy occasioned by the availability of these products has been enormous, with predictions of massive job loss and transformative change in many spheres of human activity, including commerce, entertainment, and education. Language learning too is poised to see profound changes through the availability of AI, although there is a great deal of uncertainty over what kind of future AI will bring in its wake.

A direction that will likely shape the role AI plays in language learning is that AI systems are becoming increasingly multimodal. They are supplementing their text dataset through being trained on a rich array of non-textual data including images, sounds, and video. This is already underway with the development of Google's *Gemini Ultra* and OpenAI's ongoing *ChatGPT* series (Metz, 2024). The same machine learning process that has enabled AI systems to generate convincingly texts of all kinds (essays, fiction, poetry, computer code) will make AI systems capable of working within a variety of media types. They not only will be able to generate pictures or video clips based on text prompts but also deliver multimodal output based on user-supplied images or audio. This will allow AI systems to be integrated more seamlessly into today's digital world, which has become increasingly multimodal, in large part through the ubiquity of multimedia-friendly mobile devices.

AI text generation already offers compelling opportunities for L2 practice and focused study, including conversational exchange, generating study plans, offering corrective feedback on L2 use, providing exercises on demand, and much more (Baskara, 2023; Godwin-Jones et al., 2024; Kostka & Toncelli, 2023; Lorentzen & Bonner, 2023). The major AI tools (*ChatGPT*, *Gemini*, *Bing*) have integrated voice interactions (speech in and out) into their chatbots, sometimes through mobile versions. That provides opportunities for listening comprehension and speaking practice that is much more compelling than what has been available through personal digital assistants (*Siri*, *Alexa*) or with the scripted interactions in chatbots available in commercial language learning apps (*Mondly*, *Duolingo*). Given the prominence today in social media of images and video clips (*Instagram*, *TikTok*), adding those media to

AI interactions is likely to increase learner interest, leading potentially to more interactions in the L2. The integration of visual and auditory media with verbal language reflects the reality of human communication. A semiotic perspective emphasizes how important signals beyond verbal are for meaning making (Godwin-Jones, 2023a; Klimanova & Lomicka, 2023).

The use of multimodal AI will add an expanded semiotic dimension to text generation, as one example will show. One of the useful capabilities of *ChatGPT* is to generate stories. Teachers (or learners) can instruct the AI system to create a particular kind of story (fairy tale, sci-fi adventure, animal fable, etc.) at a specific proficiency level (A2, B1) and at a particular length. AI systems can accomplish that task in a variety of languages. AI can also be prompted to create a variety of wrap-around learning activities based on the created story, for example, advanced organizers, comprehension exercises, or role-play activities. With multimodal capabilities, an AI system can insert appropriate illustrations as well as supply the option to have a story read aloud. A video version of the story or an anime comic book could also be generated. We know from SLA research that multiple representations of a story, such as in mixed media, aid in recall and retention (Plass & Jones, 2005). Dual coding theory (Paivio, 1990) and multimedia learning theory (Mayer, 2020) have shown how pedagogically effective the elaborate processing of information (going beyond a single modality) can be. Those theories are confirmed in more recent studies relying on neuroscience (Jeong & Li, 2023; Legault et al., 2019a; Li & Jeong, 2020).

While it is already possible to create an AI-based multimodal story, it requires use of several different tools or services. Multimodal capable AI will simplify the process, making it considerably more accessible to teachers and learners. Digital storytelling has been shown to be an effective learning and motivating activity for many learners (Godwin-Jones, 2012; Wu & Chen, 2020). One could well imagine students developing personal stories of their own and working within AI to make them multimodal. Storytelling apps like *Storybird* or *Scratch* could be integrated into a storytelling flow line. Those apps have been shown to foster student creativity and originality and hence provide a vital human counterpart to a process featuring AI (Peña-Acuña & Navarro-Martínez, 2023). Digital storytelling offers many options in terms of media integration (photos, graphs, drawings, animations, texts, audio clips) as well as multiple development tools/platforms. This gives learners considerable authoring freedom, providing an opportunity to become creative agents, finding and integrating artifacts of personal significance. Qi (2025; this volume) demonstrates how collaborative digital storytelling can inspire identity exploration and expression.

An illustration of the integration of AI-generated multimedia into storytelling is laid out in Hsiao et al. (this volume). Taiwanese high school students learning English developed collaborative websites based on a novel they had read (*Wonder* by R. J. Palacio, 2012). After writing their own stories connected to the novel, students used *ChatGPT* for proofreading, then turned to AI image creation tools to generate accompanying illustrations. Their illustrated books were created in the form of flipbooks, audiobooks, or videos. Contrary to the oft-expressed view that AI use poses a grave danger to human creativity, this project, featuring a collaborative process among students and AI tools, "unlocked their potential and encouraged them

to push their boundaries." That kind of learner agency through AI was available even to low achievers in English, while all students gained digital literacy through participation in the project.

The success of the *Wonder* project was facilitated by teacher guidance, an essential ingredient in effective AI use in education. Many teachers, however, are likely to be highly ambivalent about using AI. Part of the issue is the fast pace of innovation. Given the industry's pace of change, it is no easy task to follow developments. That rapid rate of change can be illustrated by the transition toward customized AI solutions. Teachers beginning to use AI will have gained experience writing prompts (to develop lesson plans, design exercises, create readings, etc.). That process is likely to lead to insights into the art of "prompt engineering," including the necessity to revise and refine instructions to achieve desired output. While that knowledge remains valid and valuable, new options for instructing AI are developing. Indeed, the process of writing prompts is undergoing a transformation. One approach is to use AI itself to create a customized prompt, as detailed in Gibbs (2024).

An alternative to general AI services that is gaining traction is to create a customized AI implementation. *ChatGPT* 4 Plus (the paid version) enables standalone "GPT" apps to be created and distributed. This allows a teacher to create an instance of the AI for a dedicated purpose, a grammar tutor, for example, or an essay evaluator. Once created, the custom GPT is given a unique URL which can be provided to target users. Such a "pedagogical AI agent," created in *ChatGPT* 4 plus, is described in Lan and Chen (2024), designed to tutor English learners in the use of ordering and transitional terms in narrating a sequence of events. The "Cool English" website has moved in that direction by creating custom chatbots in different categories that enable conversation on a variety of themes and in different scenarios, including the ability to "chat" with famous athletes, celebrities, or politicians (Chen et al., 2025; this volume). A special grammar checker for English based on *ChatGPT* is also available as a customized tool through that site. It is possible for students themselves to create a custom GPT, a process which can be both highly motivating and pedagogically effective (learning through teaching). An example is outlined in Yeh and Lan (2018) of a student-created VR project. Having students take on the role of teaching an AI system a second language (AI as a tutee) is described in Tu et al. (2025; this volume). Multimodal AI holds the promise of enabling learner creativity in modes likely familiar from social media use, building the potential for enhanced learner motivation and agency.

11.3 AI's Language Abilities

AI systems which have been trained on image analysis as well as on texts have a better understanding of physical reality (Xu et al., 2023). That understanding will be considerably enhanced through AI gaining video abilities. In fact, one of the limitations of systems like *ChatGPT* resides in its text-only training data. Generative AI relies on machine learning applied to an immense dataset of text to find patterns

and regularities that allow it to predict the next item in a text sequence. Current AI systems are not built with the goal of evaluating text through anything like a human reasoning process. Interestingly, that was a goal of early AI, which was trained through annotated text samples, intended to teach the system to understand language the way humans (or at least linguists) do, that is, as a set of rules governing syntax, morphology, semantics, and phonology, the linguistic building blocks of human language. At the same time, computer scientists fed into AI systems knowledge about how the social and physical worlds operate. That was the idea behind the famous CYC project (Lenat, 1995). One could reasonably expect such a system to have a fundamental understanding of our world and therefore to create output both reasonable and factually correct. However, this approach—teaching AI about the human language and the social world—met with only limited success. Our world and human society are much too complex to be reduced to a set of programming instructions, no matter how vast or ambitious.

A much more effective approach has been the transformer model to building LLMs, which uses machine learning through multi-layered artificial neural networks to find the next most statistically likely token (text chunk). That enables generative AI systems not only to understand language but to produce original texts which are of a quality high enough to appear to be authored by humans. In analyzing language, AI systems do not use a linguistic model (as has been traditionally the case in tutorial CALL), but rather a mathematical model, created by transforming speech (words, phrases, sentences) into mathematical entities (vectors). These statistically based systems have a remarkable ability to generate coherent discourse. However, generative AI systems do not have any real understanding of the language they produce. For AI, language is a purely statistical operation. In effect, AI systems are repackaging language based on their training data, with the result being unique and often impressive, but not creative or original in the way that humans can use language. Fiction and poetry written by *ChatGPT* are bland and uninspired compared to that of human writers (Elam, 2023; Kramsch, 2023; Woo et al., 2023). While output may be coherent and grammatically accurate, it can also contain surprising factual errors ("hallucinations") as it lacks the common sense even small children develop from having lived in the real world (Kosinski, 2023).

Humans learn and use language quite differently from generative AI. Qi (2025, this volume) reminds us that human language is not only acquired, but also experienced and lived. Children learn to use language through socialization, beginning within a family and expanding into ever larger social circles as we grow. We learn language in the context of everyday life around us, thus integrating linguistic and social learning. Studies in L1 acquisition have shown that learning occurs through interpersonal interactions, not just through observation and imitation (Kuhl et al., 2003). Social learning is a reciprocal process, requiring joint attention (Yu & Smith, 2016). Through contact with others, we acquire socially appropriate ways to use language, i.e., pragmatic competence. That includes appropriate nonverbals for the context, such as facial expressions and paralanguage. It is increasingly recognized in SLA theory that human communication is embodied and contextually embedded

(4E theory; Ellis, 2019). Gestures and gaze are essential components of how humans make meaning (Taguchi, 2021).

AI systems trained exclusively on text samples do not acquire the ability humans have to deal with language as an embodied and socially determined phenomenon. While they are able to learn aspects of pragmatic language behavior, such as politeness formulas, they have been less able to deal with pragmalinguistic interactions that might involve complex social and cultural realities (Lee & Wang, 2023; Su & Goslar, 2023). *ChatGPT* has been shown to be weak in interpreting implied messaging (Qiu et al., 2023; Ruis et al., 2023). Chen et al. (2024), on the other hand, found that in speech act performance, *ChatGPT* performed equally well to human participants. However, in that project, the AI system was fed "supplemental data" (not provided to humans) that functioned as attitudinal indicators (i.e., "write the conversation as if you feel apologetic"). *ChatGPT* could not generate appropriate responses based on its own but needed to be told how to react. The authors themselves expressed a doubt "whether AI-generated conversations can convey effectively information of speakers' subjectivity, such as their identities and stances" (p. 27).

Dynel (2023) found that "*ChatGPT* comes across as overly polite, employing multiple politeness formulae and all manner of strategies to avoid any kind of face-threat to the interlocutor or any other individuals" (p. 117). That *ChatGPT* goes overboard in its cooperative stance, exaggerating the well-known Gricean maxim (Cooperative Principle; Grice, 1989), may be a result of human reinforcement training or other instructions/data provided to the system. The fact that AI systems like *ChatGPT* react in pre-determined ways, such as being overly polite, undermines their ability to adjust to personal differences in respect to pragmatic norms. While individuals may be aware of expected cultural norms for speech acts (degree of directness in requests, accepted ways to apologize) or for conversational exchanges (rules for turn-taking; avoidance of periods of silence), they may, for personal reasons, opt to violate expectations (see examples of resistance and divergence in García-Pastor, 2020; McConachy, 2019). The social use of language is dynamic, contingent, and personal. It can also be unpredictable and variable, so that in fact, there is a wide range in allowable pragmatic behaviors (Bardovi-Harlig, 2020). Children are more effective language learners largely because they acquire language, not from rules, but in socially interactive contexts (Li & Jeong, 2020). Studies in brain science show that grounding second language learning in social interaction can enhance retention (Jeong et al., 2010), a finding that accords with sociocultural theory for second language acquisition (SLA; Lantolf et al., 2014). Pragmatic competence has been shown to develop slowly (Barattieri di San Pietro et al., 2023) through repeated social encounters and is not easily reducible to rules and set patterns.

AI systems have a problem with pragmatics because they lack "*direct sensorimotor grounding* to connect its words to their referents and its propositions to their meanings" (Harnad, 2024; emphasis in original). Multimodal AI is likely to improve significantly the ability of systems to interact with humans in ways that incorporate nonverbals and pragmatic language use. There is already some indication that generative AI systems are capable of learning some socially appropriate

ways to communicate. That includes speculation that such systems may spontaneously develop a "theory of mind," the ability that humans possess to perceive and take into consideration the mental and emotional state of our interlocutor (Kosinski, 2023). Barattieri di San Pietro et al. (2023) suggest that "part of what we call pragmatic competence might be coded in the regularities of the language use and that hence, distributional models are well-equipped to capture and exploit such regularities" (p. 389). Researchers have begun probing into the abilities of such systems to incorporate an understanding of human emotions (Mok, 2023; Potamianos, 2023). The ability of generative AI systems to learn is, in fact, one of their greatest strengths. The creators of *ChatGPT* were surprised by some of the emergent capabilities of that system, its ability to do programming or machine translation, functionalities they had not anticipated (Kosinski, 2023).

The unanticipated capabilities of generative AI illustrate how radically different LLM-based systems are from traditional software programs. In classic software development, programmers write each line of code, understanding the process from beginning to end, with outcomes known and predictable. In contrast, generative AI development is largely independent of its creators, although "reinforcement training with human feedback" is used to fine-tune the system and seek to eliminate errors. The inside workings of AI remain largely a black box (see O'Gieblyn, 2023). Becoming aware of that fact is fundamental to developing an understanding of generative AI and therefore a step toward critical AI literacy. A lack of transparency is likely to increase as AI systems develop more effective self-learning techniques, rendering them autonomous—and therefore unpredictable—learners (Godwin-Jones, 2024b). So far, calls for "explainable AI" (Ali et al., 2023) have not yielded impressive results.

11.4 Toward a Metaverse

The ability of AI systems to incorporate more completely the full panoply of human mechanisms for communication has important considerations for an emerging metaverse and the development of extended reality for language learning. Until now, the fledgling metaverse has been far from the imagined merger of the digital and physical worlds. On the one hand, real-world data has become readily available through the multiple cameras, sensors, and stored online media that track our daily activities. Indeed, privacy concerns abound given widespread monitoring and surveilling. To what extent tracking data will be integrated into a metaverse is an open question; guaranteeing privacy and preventing nefarious uses of personal data will be challenging issues. The potential exists, however, that with appropriate safeguards and user consent, individual digital profiles can be created that are more comprehensive than the minimal information now retained in systems such as voice assistants or user accounts. That information would include not just general information about the user (subscribed services, for example), but deduced info such as hobbies and personal interests, valuable in the creation of virtual representations. Pioneers of *lifelogging* (extensive individual daily tracking) have embraced that vision (Dingler et al., 2021).

While real-world data is abundant, extensive virtual 3D worlds on the other hand, a crucial component of the metaverse, are limited. Creating immersive, realistic, and dynamic backdrops such as those available in high-end gaming has been an expensive and time-consuming process, involving designers, graphic artists, animators, and programmers. Mirrored versions of existing sites, created by stitching together multiple 360-degree video clips, are also arduous. Yet a metaverse, conceived as a location for entertaining, business transactions, and socializing, will need to combine fantasy scenarios with simulated real-world venues. That rich virtual world will need to be continually available to unlimited numbers of users through a variety of interconnected devices and platforms, with the degree of access and interactivity determined by the device's capabilities. In the conception of the metaverse, the virtual world is linked closely to the real world, so that actions in the virtual space affect the outside world and vice versa. The 3D shopping and marketing platform TheMall (https://themall.io/) is an illustration of that merged space.

Creating such a digital ecosystem is a tall order to pull off technically and creatively, but the arrival of multimedia AI promises a viable alternative for building the 3D content needed (Huynh-The et al., 2023): "Generative AI will elevate the quantity of content in the metaverse to a new level and drive the resurgence of industries such as virtual reality (VR) and augmented reality (AR)" (Lv, 2023, p. 208). AI systems will be able to build autonomously both imaginary worlds and replicas of real-world settings. Through emerging video creation capabilities, these virtual worlds—and the avatars that inhabit them—can become responsive and dynamic, reacting in real time to user actions. Moreover, generative AI lowers dramatically the technical threshold, as media development no longer requires traditional programming, but rather involves writing prompts to provide detailed instructions. Companies like Agora World and METAVRSE have created no-code virtual world development platforms.

It has been shown that complex problems can be addressed in AI through "chain-of-thought" prompting (Malach, 2023; Wei et al., 2022). Additionally, AI systems have begun to accept much longer prompts and also to enable documents with detailed instructions to be uploaded and added to the system's database. Multiple AI systems can be combined to create a custom pipeline for a hybrid system (Chen & Yeh, 2025, this volume). The additional uploaded data or linked systems could easily be related to specific aspects of language or culture. While AI systems possess multilingual capabilities, incorporating specific pedagogical and linguistic data could enhance the system's usefulness for language learners (see Godwin-Jones, 2021). The incidental language learning potential in a multilingual metaverse is already considerable, but targeted learning could be made possible as well. In that way, a mix of implicit and explicit learning can be achieved, a combination often seen in SLA theory as optimal (Ellis, 2009; Hulstijn, 2005).

Virtual reality, the essential gateway into the metaverse, has been slow to gain traction in everyday activities such as shopping or working. For the most part, VR has been mostly a haven for gamers (Cureton, 2023). Indeed, pundits have pointed to the expansion of gaming platforms such as *Fortnite* or *Roblox* as moving in the direction of a burgeoning metaverse (Park, 2020). The availability of those platforms on a

great variety of devices and systems, from phones to gaming consoles, points to the ubiquity of access needed for this vision. Required as well for a metaverse are other elements already available within *Fortnite* or *Roblox*: the ability to have both planned and spontaneous events, to offer a variety of gaming and communication options, and to have its own economic system (such as gaming currency) and, importantly, to enable users to carve out within that environment their own space, offering, for example, user-created games.

The limited use of VR outside of gaming—and therefore the slow emergence of anything resembling a "metaverse"—is due to factors that have plagued immersive VR since its inception: practicality, comfort, and compatibility. Availability is constrained by the high cost of VR headsets. The complete isolation from user surroundings tends to suggest self-absorption, if not anti-social behavior. Added to that is the physical discomfort of prolonged use, as well as the disorientation that can lead to feelings of nausea. Headset manufacturers are working to reduce costs and weight, and it seems likely that the trend toward ever smaller digital components will enable that development. At the same time, mixed reality headsets—like Apple's *Vision Pro* or Microsoft's *Hololens*—enable more flexible use including integrated augmented reality (AR), which allows the user to see and interact with the immediate environment. For the vision of a metaverse to become reality, VR and AR devices will need to become ubiquitous and interoperable, a development not yet on the horizon.

It has been easier for AR to function across devices than it is for VR, where apps may require a specific headset brand and version. For AR, apps are normally available in both iOS and Android platforms and simply require a camera on the mobile device. In contrast to VR, AR integrates the user into the immediate surroundings, so that apps can use GPS and local landmarks to create location-aware experiences, such as language/culture tours (Pegrum, 2021) or collaborative games (Holden & Sykes, 2011; Thorne, 2013). The embedded, situated learning enabled through AR is likely to be motivating and memorable (Aldossari & Alsuhaibani, 2021). An advantage of AR is the likelihood that all students will possess the necessary hardware (i.e., a smartphone), thus making the technology available both in the classroom (maker-based AR) or in extramural contexts (Godwin-Jones, 2024a).

Pegrum (2021) suggests that in the near future, AR will be accessed through smart-glasses, such as those demonstrated by Google in 2024 or available from Ray Ban or Amazon. That will enhance AR functionality, but at the cost of ubiquitous availability. The ideal vehicle for extended reality may be mixed reality glasses or a headset that is lightweight and flexible, usable for both AR and VR. Apple's *Vision Pro* points in that direction, but—in its first iteration—at a prohibitive cost and with unattractive bulkiness. It offers, however, intriguing innovations that may point the way for future devices in a metaverse. Rather than requiring hand controllers, Apple has integrated down-facing cameras that capture user gestures and hand movements. Device control is enacted through body movement, voice commands, and eye tracking. The default mode of *Vision Pro* is AR; a digital crown enables immersive mode for VR use.

While mixed reality devices supply the normal entryway into the metaverse, the growth of the *Internet of Things* will supply additional means for input and output. Smart devices such as home speakers, car systems, and wearables are incorporating

AI, often by way of commonly used voice assistants, integrated now with AI. Dedicated AI devices have become available as well, such as the Rabbit AI (smartphone form factor), the Humane AI Pin (attached to clothing), or the Rewind Pendant (worn around the neck). The *ambient intelligence* surrounding us (Sadri, 2011) will provide digital services (most often voice activated) wherever we happen to be, clearly an intriguing future for language learning, in which our everyday life can be integrated into learning scenarios, making SLA personal and ubiquitous. We know from studies on L1 and L2 learning that language encountered in familiar, real-world settings—and especially if connected to concrete everyday objects—is more likely to be learned long-term than the same material practiced in a classroom setting (Chen & Yeh, 2025, this volume; Jeong & Li, 2023).

11.5 VR in Language Learning

Extended reality use for language learning has seen a history—like AI—of mixed success and failure (Godwin-Jones, 2023b). Neither AR nor VR is used widely today in language learning, either in formal instruction or through informal autonomous learning (Berns & Reyes-Sánchez, 2021). Dhimolea et al. (2022) conclude from their analysis of VR for SLA that studies of VR show "potential rather than evidence" of usefulness in instructed language learning (p. 820). In prior incarnations, however, VR has proven to be both popular and effective, although in a technological implementation very different from today's immersive VR. The first wave of early VR consisted of text-based adventure and simulations platforms, principally multiuser, object-oriented environments (MOOs) like LambdaMOO in the 1990s, which provided an exploratory, collaborative L2 experience (Lin & Lan, 2015; Schwienhorst, 2002). The 2nd wave, desktop VR, mainly, *Second Life* (2003), continued this trend with an emphasis on peer interactions, experimentation, and the development of learner autonomy. Hubbard (2019) follows the rapid rise (and decline) of *Second Life* in language learning. For a discussion of low-immersion VR (i.e., using a 2D computer screen and mouse), see Sykes et al. (2008). *Second Life* continues today to be a viable platform for language learning, especially when used in conjunction with advanced voice technologies (see Grant, 2025, this volume).

For the most part, immersive VR has gone in a different direction from earlier VR. Immersive VR—using 360° surround headsets and hand controllers—is the dominant version of the technology today (Berns & Reyes-Sánchez, 2021). This wave of VR dates from the 2010s, when lower cost headsets began to be available. The inexpensive Google Cardboard, released in 2015, created a spike of interest (Dhimolea et al., 2022); its functionality, however, is quite limited in comparison to full-fledged VR headsets. If earlier implementations of virtual worlds in language learning emphasized collaboration and exploration, developments in immersive VR have instead involved mostly multimodal vocabulary learning and fixed role-play activities. Learning scenarios present mostly tourist-like experiences (taking a taxi,

ordering at a restaurant) with the user role limited to responding to closed, multiple choice options. Kronenberg and Poole (2022) comment:

> One noticeable trend in the application of VR for language education is the creation and exploration of VR environments that are incredibly limited in scope and application. Many of these VR environments are used in research studies as highly immersive flashcard systems (e.g., Legault et al., 2019b). Such studies send the message that VR systems are no more than a novel tool to be used to inspire learners to spend a few more minutes looking over vocabulary words. (p. 3)

Other researchers have echoed this view (Berns & Reyes-Sánchez, 2021; Nicolaidou et al., 2021).

While apps which provide immersive vocabulary practice and structured role-play activities can play a useful role in language learning, especially at the novice level, they barely tap into the immense potential of immersive VR. This is a medium which can supply rich, multimodal social and cultural interactions in an environment which can faithfully mirror existing settings (a Paris neighborhood; a market in Mexico City) or create entirely new worlds (science fiction; fantasy adventure). Multimodal simulations offer learners the opportunity to use their L2 through taking on different identities and using the L2 discourse aligned with those personas (Klimanova, 2021). Global simulations in digital environments have shown how identity positions in such *mirror worlds* (Kelly, 2019) can go beyond learners as tourists (Michelson & Petit, 2017; Mills et al., 2020). Mirror worlds in VR environments can supply multiple layers of reality, including past representations. VR apps, such as *Immerse* or *vTime,* offer functions that enable learners in a simulated 3D world to interact with each other, allowing for collaborative learning. A VR app such as *WonderVR* allows for creating simulated environments through capturing a series of video clips using a 360° camera. VR apps for language learning will inevitably offer greater opportunities for more spontaneous and variable interactions through the integration of generative AI, a development already underway (see Chun et al., 2025; this volume).

Depending on the implementation and the equipment used, VR adds an additional, kinesthetic dimension to immersive interactions, tracking body movements. That could be used in language learning to experience virtually cultural practices tied to a particular context. A language learning game, *Crystallize*, was adapted to upper body movements in a project that enabled Japanese learners to experience different bowing scenarios and to receive feedback (Culbertson et al., 2016). The incorporation of gestures was a central aspect of a VR-enabled exchange of English learners, using the social VR platform, *Spatial* (Chen & Sevilla-Pavón, 2023). Room-scale immersive environments supply more precise tracking of user movements, as in the vocabulary learning project described in Vásquez et al. (2018). Innovative uses of VR do not necessarily require creation of virtual environments from scratch, which is likely unpractical for most language teachers. Off-the-shelf VR games offer opportunities for language learning, especially if thoughtfully integrated into a task-based approach with related language learning activities. Kronenberg and Poole (2022) provide an illuminating example of using a commercial game, *Keep Talking and Nobody Explodes*, for collaborative language learning.

Integrating gaming into an instructed language environment offers students a welcome alternative to normal learning activities and can be highly motivating (Reinhardt & Thorne, 2020). Also motivating is providing learners opportunities to function not just as users of VR but as creators. In the "City in the sky" project, learners used a variety of simple VR tools to engage in highly personalized, multimodal storytelling (Chun et al., 2025, this volume; Karimi et al., 2023). That project, in contrast to most VR implementations, which are short term, extended over a longer time frame, allowing for extensive personal exploration and peer collaboration. As expressed by Peña-Acuña and Navarro-Martínez (2023), "creativity unfolds over time, influenced by social interaction and the environment" (p. 2). The collaborative and multimodal storytelling in this project allowed each participant to leverage individual interests and backgrounds, contributing in different content areas and modalities, thus exemplifying a "multiliteracies" framework (New London Group, 1996). The longer project length also provided an opportunity for discussion and reflection, an essential attribute of successful learning implementations of VR or of gaming (Blyth, 2018; Reinhardt & Thorne, 2020). Another characteristic of successful VR projects is that they provide sufficient learner orientation and training before actual VR use (Lloyd et al., 2017). This is especially important for learners new to VR. Without preparation, there is the potential for cognitive overload due to the combination of L2 stress and navigation in an unfamiliar technology (Chun et al., 2025; this volume).

With the arriving integration of generative AI into VR, new, rich learning opportunities abound. Scripted interactions in simulations and role-play activities can now be replaced with dynamic, on-the-fly encounters. If those interactions occur in an immersive environment, they can take advantage of the multimodal capabilities to provide engaging visual and auditory elements. *ChatGPT* and other AI systems can easily take on different identities, with content and language imitating closely those of an existing person or of a persona designed by the user or teacher ("in this conversation take on the role of a 25-year-old Parisian hairdresser"). Identities could include historical persons (Albert Einstein, Martin Luther King) or fictional characters (Harry Potter, Indiana Jones). Role-play scenarios could be wide-ranging with detailed prompts setting up the exchange. AI itself could be used to suggest a storyline.

A multimodal, interactive story involving L2 learners and a real or fictionalized character could be customized both in terms of personal learner interest ("I want to sing along with Taylor Swift") and learner's L2 proficiency level. AI could supply learning materials of a variety of types. In an instructed learning environment, the setup for such an activity could include instructions to integrate specified language constructions or vocabulary that align with course content. Of course, role-play scenarios that involve actual human beings need to be treated with caution (Godwin-Jones, 2024b). The deep fakes that can make VR games compelling may violate privacy or intellectual property rights. The actors strike in Hollywood in 2023 occurred in large part due to the ability of AI to replicate actors' voices and likenesses.

11.6 AI and VR: The Need for an Ecological Framing

In AI-based VR, smart avatars will be capable of displaying facial expressions appropriate to the context; in effect, avatars will become AI agents. AI robots can go further, with the ability to incorporate gestures and body language. In fact, tangible objects in the immediate environment could become learning opportunities (see Chen & Yeh, 2025, this volume). As AI becomes integrated into ever more devices, language learning will no longer be bound by place. That fact should lead us away from viewing AI simply as a tool but rather as one component in a larger, holistic learning environment (Qi, 2025, this volume). As is the case for any digital tool or service, AI cannot be analyzed in isolation, but needs to be viewed in the context of use. This is particularly true for disruptive technologies which change not only educational practices but our very way of life. We have seen the social and cultural transformation ushered in first with the World Wide Web and browsers in the 1990s and most recently with the smartphone, dating from 2007. Our everyday world today is hardly imaginable without the availability of mobile apps we rely on for socializing, navigating, communicating, and more: "The reliance on apps and functions has led to a level of dependence and personal intimacy new in the human–machine relationship, with the devices offering 'extensions of human cognition, senses, and memory' (Moreno & Traxler, 2016, p. 78). Smartphones, for many of us, have indeed become an extension of ourselves—something like a digital appendage" (Godwin-Jones, 2017, p. 4). *ChatGPT*, and the AI wave it created, heralds a transformation similar to that of the smartphone, with changes affecting all aspects of society, including work environments, education at all levels, and many aspects of leisure activities including gaming and social media.

Helpful in understanding more holistically the impact of AI on our lives and in particular its role in language learning is to examine its use through ecological frameworks, which explore the complex relationship between users and the environment. One of those used increasingly in applied linguistics is sociomaterialism (Qi, 2025, this volume; Thorne et al., 2021). Rather than examining an artifact or a tool in isolation, a sociomaterial perspective looks at its context of use, particularly the respective social and cultural dynamics (Guerrettaz et al., 2021). It offers a way to disentangle the "assemblages" of interactions between humans and non-humans (Sørensen, 2013). Studies of the use of *ChatGPT* for L2 learners have demonstrated that interactions with users are neither simple nor one way, but rather typically involve a continuous back-and-forth series of iterative prompts and text outputs (Tseng & Warschauer, 2023; Warschauer et al., 2023).

That dynamic is illustrated in a case study of a Chinese graduate student in California, "Kailing," using *ChatGPT* in academic English writing (Jacob et al., 2023). She uses the tool in various stages of the writing process: in brainstorming ideas, doing basic research, generating sections of a first draft, and assisting in revisions. She revises prompts in an iterative process, going back repeatedly to *ChatGPT* for new output, allowing her to fine-tune texts to her need. Interestingly, she at times rejects AI suggestions, as they do not mesh with her own voice and style, which

she finds important to retain. In that way, there is a distributed or shared agency between AI and the user, not a simple tool use (Godwin-Jones, 2024b). Another example of an evolving partnership between AI and learners/teachers is evident in the "team-teaching" approach to deploying an adaptive, intelligent tutor outlined in Lan and Chen (2024, p. 2). While the English teacher designs the learning interactions, continuously tests system revisions, and mentors students pedagogically and emotionally, the AI tutor is put in charge of "controlling learning flow, personalized content delivery, instant feedback, and performance tracking" (Lan & Chen, 2024, p. 3). The shared responsibilities and distributed roles that characterize these uses of AI demonstrate the potential complexity of the relationship between AI and users.

A sociomaterial perspective is particularly helpful in understanding the dynamics of AI-based VR (Godwin-Jones, 2023b), in that interactions potentially go beyond verbal language, with digital assets such as 3D images, animation, or non-playing avatars creating a complex interplay of human *presence* (and *copresence*) with non-human entities. Another ecological framework which emphasizes relationality is complexity theory or complex dynamic systems theory (Godwin-Jones, 2018). Originating in the biological sciences and related to chaos theory (Larsen-Freeman, 1997), complexity theory addresses how unexpected, emergent outcomes may arise from small changes in initial conditions or in ongoing interactions with people, artifacts, or situations. For language learning, it emphasizes the variability of outcomes, even in a homogeneous classroom setting, as learners each have a unique learning trajectory depending on previous experiences with the target language and personal characteristics.

Adding AI into the language learning mix (itself a complex dynamic system) brings an additional element of uncertainty and unpredictability. A fundamental tenet of complexity theory is that of *emergence*, the coming of something new and unanticipated that arises from the dynamic interplay of elements in a system (Larsen-Freeman, 2022). Kailing's completed academic writing (Jacob et al., 2023), a unique hybridized creation from her collaboration with *ChatGPT*, provides a compelling example of emergence. That experience also demonstrates the lack of a central controlling authority typical of complex dynamic systems. From that perspective, agency is shared and relational, enacted contingently as the components of the systems interact. According to Sørenson (2013): "The relationship between human and non-human does not need to be seen as a clear-cut dichotomy or as a stable or permanent continuity, but can also unfold as sliding positions in a spectrum of changeable possibilities" (p. 10). Relationships need not be symmetrical among humans and non-human actors, with initiative and control shifting depending on need and context.

With all the possible learning resources available today, learner pathways are more variable than ever. That should lead to an acceptance of differentiated outcomes and the need to examine retroactively the mix of resources leading to successful (or unsuccessful) language learning. That emphasis follows trends in SLA theory that stress the importance of considering individual learners holistically as complex human beings, not simply as "students" (Benson, 2017; Ortega, 2017). Learners should be optimally treated, as "developing selves" (Qi, 2025, this volume), with multiple versions of their identities emerging as they live and learn, shaping their

present lives and anticipating their future selves. AI-powered VR is likely to supply an ideal vehicle for identity exploration.

In the emerging AI-empowered metaverse, originality, authenticity, and attribution become complicated, as shifting contributions from human and AI blur concepts of sole authorship (Elam, 2023). For both L2 learners and teachers, navigating these spaces will be difficult, as traditional assessment mechanisms will become increasingly problematic (Cope et al., 2021). The ability of L2 learners to use highly effective text generation tools and reliable machine translation may mean that linguistic/grammatical accuracy "can no longer be viewed as a synonym of learning and excellence" (Klekovkina & Denié-Higney, 2022, p. 107). It may be that as AI tools proliferate, less instructional attention will be paid to surface-level errors (spelling, grammar) and more to higher level language issues (originality, organization, personal voice). Learners may be asked to report their use of AI as well as to reflect on its usefulness. Studies on the use of machine translation may be helpful in that regard (Hellmich & Vinall, 2021; O'Neill, 2019; Vinall & Hellmich, 2022). These studies outline a number of approaches to assignments that make meaningful and ethical use of machine translation (post-editing, comparing human and machine drafts, investigating translanguaging).

In any case, a comprehensive L2 language and writing program will supplement AI-enabled authoring with activities that emphasize social aspects of language use and target directly human-to-human communication. That could include activities such as peer editing or the use of non-disposable or *renewable assignments*, i.e., those that go beyond classroom-based artificiality to real-world integration (Blyth, 2023; Ryan & Kautzman, 2022). Collaborative activities which involve negotiation of meaning and the use of strategic/interactional competency are particularly important. Virtual exchange provides an ideal vehicle for that purpose (Godwin-Jones, 2019; Qi, 2025; this volume). Interestingly, virtual exchange recently has seen promising examples of VR integration (Baralt et al., 2022; Yeh et al., 2022).

11.7 Conclusion

The original metaverse appeared in Neal Stephenson's *Snow Crash* (1992). In that novel, as well as in the similar vision of "cyberspace" in William Gibson's *Neuromancer* (1984), this is a dystopian vision, with continuing human consciousness after death portrayed in nightmarish terms. It remains to be seen what kind of environment for human well-being a metaverse might offer. The prominent role that AI is likely to play in building and maintaining the metaverse raises a host of concerns related to human well-being. AI scientists themselves see a possible threat to the future of humanity through runaway or rogue AI (Roose, 2023). Beyond that doomsday scenario, the process used to create LLMs holds many ethical issues, including intellectual property rights, fair wages for human AI workers, inherent biases in data collection, and the environmental costs of running immense racks of power-hungry supercomputers (Godwin-Jones, 2024b). There are fairness, equity, and accessibility

concerns as well in that most AI companies are private, with a mission to make a profit, not to benefit humanity.

The most capable models of AI will inevitably come at a cost that will be prohibitive for many. That is a serious issue for education, where the digital divide is already strongly in evidence (Warschauer, 2004). Possible alternatives to commercial AI systems are open-source LLMs, which have become more widely available. In fact, some of the big players in AI, Meta and Google, have themselves made at least some of their AI models freely available. More positive public views of AI—and of a developing metaverse—would be helped by companies agreeing to make their leading-edge products more accessible. Government regulation of AI development does not look promising, but one could hope that governments might have enough influence to push companies to commit to a "moral AI" that learns to intrinsically care about human lives (Qureshi, 2023).

In terms of language learning, it is clear that the powerful multimodal and simulation capabilities of AI-based VR will provide users new, engaging learning opportunities. The ever-growing realism possible in virtual worlds is likely to foster informal language learning that simulates how we learn our first languages. Indeed, studies in brain science have pointed to that potential in VR (Li & Jeong, 2020; Ma & Yan, 2022). The personalization possible through AI, along with the attractive multimodality of VR, is likely to be highly motivating learning factors. Studies in informal language learning through streaming video or social media have shown how incidental learning can take place as a consequence of engaging in entertainment or socialization (Cole & Vanderplank, 2016; Godwin-Jones, 2018; Sockett, 2014). While general AI tools will offer a default metaverse environment, it seems likely that hybrid versions will be used as well, allowing for specific language learning frameworks to be incorporated into the metaverse (see Chen & Yeh, 2025, this volume). This is the future of intelligent, contextualized learning (Lan, 2025, this volume) that has long been the dream of CALL.

In that vision of the future, AI and VR will need to be viewed as powerful aids to learning, but not adequate to replace human-centric activities (see Aslam & Li, 2025; this volume). The role of language teacher will continue to be essential, as a guiding and mentoring presence. That includes active guidance in using advanced technology. Language learning can be an emotionally fraught experience. In contrast to a human teacher, AI will not provide the caring presence and emotional support for learners as developing selves. That support may be physically present in the classroom or available virtually. The relationship dynamics in an AI world are bound to become increasingly complex, with shifting roles for learners, instructors, and AI. It may be helpful in that regard to look to indigenous cultures that value highly both human-to-human relationships and non-human elements in nature and the environment (Qi, 2025, this volume). Interestingly, postcolonial studies have begun to look at AI from such a perspective (Adams, 2021; Mohamed et al., 2020).

The ontologies of non-duality and relationality that figure predominantly in many non-Western societies "treat the Earth as made up of entities that are not hierarchical or binary. That is, features such as human/non-human, sentients/objects, and mind/body are all considered equally agentive" (Canagarajah, 2023, p. 287). The

process of *becoming* is seen from this perspective as common to all entities; constant change and development through encounters with others, human and non-human, is inevitable. That process "involves non-linear trajectories, unpredictable changes, and indecipherability" (Canagarajah, 2023, p. 288). This indigenous perspective reinforces insights from sociomaterialism and complexity theory that "simple dualism between humans and non-humans are not helpful in a complex world where all things are interconnected" (Larsen-Freeman, 2019, p. 69). That interconnected perspective should inform our understanding of the role of AI in society. Rather than seeing AI use as an individual phenomenon, it needs to be considered socially. Ethics is a social phenomenon; it falls to us individually and collectively to see that AI is used responsibly.

Given that complex dynamic, teacher training in an AI world should "emphasize recognition and responsiveness over controlled planning" (Guerrettaz et al., 2021, p. 17). Teachers should not assume linear learning trajectories or universal learning pathways. Outcomes will vary according. Increasingly, universally applicable teaching methods will become inadequate. Instead, a *relational pedagogy* (Kern, 2015, 2018) emphasizes adjusting goals and methods to account for local and situational variables. The emphasis should be on the dynamic relationships among learning resources and learners. In a world merging AI and VR, those relationships are likely to become increasingly complex. Given that scenario, a *human* ecological perspective (Levine, 2020) can be helpful, an orientation that accepts non-human actors as important contributors to communication but stresses the social essentialness of human speech. That aligns with the Māori concept of *Whakawhanaungatanga*, as discussed in Qi (2025, this volume), which stresses the primacy of human relationships in our lives, as it references how we establish, maintain, and strengthen relationships.

A growing expectation for language teachers will be that they become informed consumers of AI, thus being in a position to advise and mentor students appropriately. As mixed reality options proliferate, that technology too should be on teachers' radar, particularly as its integration with AI makes learning opportunities more attractive and abundant. The future cannot be predicted, but it does seem highly likely that an emerging metaverse will provide new opportunities for language learning, both for autonomous learners and in formal instructional settings. Both learners and teachers will need to develop and maintain a level of digital literacy that enables them to use advanced technologies ethically and effectively.

References

Adams, R. (2021). Can artificial intelligence be decolonized?. *Interdisciplinary Science Reviews, 46*(1-2), 176–197. https://doi.org/10.1080/03080188.2020.1840225

Aldossari, S., & Alsuhaibani, Z. (2021). Using augmented reality in language classrooms: The case of EFL elementary students. *Advances in Language and Literary Studies, 12*(6), 1–8. https://files.eric.ed.gov/fulltext/EJ1339045.pdf

Ali, S., Abuhmed, T., El-Sappagh, S., Muhammad, K., Alonso-Moral, J. M., Confalonieri, R., Guidotti, R., Del Ser, J., Díaz-Rodríguez, N., & Herrera, F. (2023). Explainable artificial intelligence (XAI): What we know and what is left to attain trustworthy artificial intelligence. *Information Fusion, 99*, Article 101805. https://doi.org/10.1016/j.inffus.2023.101805

Aslam, M. S., & Li, Y.-H. (2025). Revolutionizing language learning: The convergence of AI and the metaverse in educational paradigms. *AI-mediated language learning in the metaverse era* (pp. XX–XX). Springer.

Baralt, M., Doscher, S., Boukerrou, L., Bogosian, B., Elmeligi, W., Hdouch, Y., Istifan, J., Nemouchi, A., Khachatryan, T., Elsakka, N., Arana, F., Cobos-Solis, J., Perez, G., Mouchane, S.-E., & Vassigh, S. (2022). Virtual Tabadul: Creating language-learning community through virtual reality. *Journal of International Students*, *12*(S3), 168–188. https://doi.org/10.32674/jis.v12is3.4638

Bardovi-Harlig, K. (2020). Pedagogical linguistics: A view from L2 pragmatics. *Pedagogical Linguistics, 1*(1), 44–65. https://doi.org/10.1075/pl.19013.bar

Baskara, R. (2023). Exploring the implications of ChatGPT for language learning in higher education. *Indonesian Journal of English Language Teaching and Applied Linguistics, 7*(2), 343–358. https://files.eric.ed.gov/fulltext/EJ1391490.pdf

Benson, P. (2017). Language learning beyond the classroom: Access all areas. *Studies in Self-Access Learning Journal, 8*(2), 135–146. https://doi.org/10.37237/080206

Berns, A., & Reyes-Sánchez, S. (2021). A review of virtual reality-based language learning Apps. *RIED. Revista Iberoamericana de Educación a Distancia, 24*(1), 159–177. https://doi.org/10.5944/ried.24.1.27486

Blyth, C. (2018). Immersive technologies and language learning. *Foreign Language Annals, 51*(1), 225–232. https://doi.org/10.1111/flan.12327

Blyth, C. (2023, June 8). *Exploring the affordances of AI tools for L2 creative writing [Conference presentation]*. CALICO Annual Conference, Minneapolis.

Canagarajah, S. (2021). Rethinking mobility and language: From the Global South. *The Modern Language Journal, 105*(2), 570–582. https://doi.org/10.1111/modl.12726

Canagarajah, S. (2023). Decolonization as pedagogy: A praxis of 'becoming' in ELT. *ELT Journal, 3*, 283–293. https://doi.org/10.1093/elt/ccad017

Chen, H.-I., & Sevilla-Pavón, A. (2023). Negotiation of meaning via virtual exchange in immersive virtual reality environments. *Language Learning & Technology, 27*(2), 118–154. https://hdl.handle.net/10125/73506

Chen, H.-J., Hsieh, F.-C., & Yen, Y.-C. (2025). Applying artificial intelligence technologies in second language learning. *AI-mediated language learning in the metaverse era* (pp. XX–XX). Springer.

Chen, X., Li, J., & Ye, Y. (2024). A feasibility study for the application of AI-generated conversations in pragmatic analysis. *Journal of Pragmatics, 223*, 14–30. https://doi.org/10.1016/j.pragma.2024.01.003

Chen, N.-S. & Yeh, H.-C. (2025). Emulating real world environment for contextual language learning using educational robot and IoT-enabled tangible objects. *AI-mediated language learning in the metaverse era* (pp. XX–XX). Springer.

Chun, D., Kaplan-Rakowski, R., Meyr, J., Ovsiannikova, U., Thrasher, T., & Yuan, Y. (2025). AI-mediated high-immersion virtual reality for language learning. In *AI-mediated language learning in the metaverse era* (pp. XX–XX). Springer.

Cole, J., & Vanderplank, R. (2016). Comparing autonomous and class-based learners in Brazil: Evidence for the present-day advantages of informal, out-of-class learning. *System, 61*, 31–42. https://doi.org/10.1016/j.system.2016.07.007

Cope, B., Kalantzis, M., & Searsmith, D. (2021). Artificial intelligence for education: Knowledge and its assessment in AI-enabled learning ecologies. *Educational Philosophy and Theory, 53*(12), 1229–1245. https://doi.org/10.1080/00131857.2020.1728732

Culbertson, G., Andersen, E., White, W., Zhang, D., & Jung, M. (2016). Crystallize: An immersive, collaborative game for second language learning. In P. Bjørn & J. Konstan (Eds.), *Proceedings*

of the 19th ACM conference on computer-supported cooperative work & social computing (pp. 636–647). The Association for Computing Machinery, Inc. https://doi.org/10.1145/2818048.2820020

Cureton, D. (2023, April 5). Is Gaming in the Metaverse the Next Frontier? *XR Today*. https://www.xrtoday.com/virtual-reality/gaming-in-the-metaverse-the-next-frontier/

Dhimolea, T. K., Kaplan-Rakowski, R., & Lin, L. (2022). A systematic review of research on high-immersion virtual reality for language learning. *TechTrends, 66*(5), 810–824. https://doi.org/10.1007/s11528-022-00717-w

di San, B., Pietro, C., Frau, F., Mangiaterra, V., & Bambini, V. (2023). The pragmatic profile of ChatGPT: Assessing the communicative skills of a conversational agent. *Sistemi Intelligenti, 35*(2), 379–399. https://doi.org/10.1422/108136

Dingler, T., Agroudy, P. E., Rzayev, R., Lischke, L., Machulla, T., & Schmidt, A. (2021). Memory augmentation through lifelogging: Opportunities and challenges. In T. Dingler & E. Niforatos (Eds.), *Technology-augmented perception and cognition* (pp. 47–69). Springer. https://doi.org/10.1007/978-3-030-30457-7_3

Dynel, M. (2023). Lessons in linguistics with ChatGPT: Metapragmatics, metacommunication, metadiscourse and metalanguage in human-AI interactions. *Language & Communication, 93*, 107–124. https://doi.org/10.1016/j.langcom.2023.09.002

Elam, M. (2023). Poetry will not optimize; or, what is literature to AI? *American Literature, 95*(2), 281–303. https://doi.org/10.1215/00029831-10575077

Ellis, N. C. (2019). Essentials of a theory of language cognition. *The Modern Language Journal, 103*(S1), 39–60. https://doi.org/10.1111/modl.12532

Ellis, R. (2009). *Implicit and explicit knowledge in second language learning, testing and teaching* (Vol. 42). Multilingual Matters. https://doi.org/10.21832/9781847691767-003

García-Pastor, M. D. (2020). Researching identity and L2 pragmatics in digital stories. *CALICO Journal, 37*(1), 46–65. https://doi.org/10.2307/27113780

Gibbs, J. (2024, January 14). *Forget prompt engineering, ChatGPT can write perfect prompts for you*. Medium. https://medium.com/@jordan_gibbs/forget-prompt-engineering-chatgpt-can-write-perfect-prompts-for-you-6ad21c4cfa99

Gibson, W. (1984). *Neuroromancer*. Ace.

Godwin-Jones, R. (2012). Digital video revisited: Storytelling, conferencing, remixing. *Language Learning & Technology, 16*(1), 1–9.

Godwin-Jones, R. (2017). Smartphones and language learning. *Language Learning & Technology, 21*(2), 3–17. http://llt.msu.edu/issues/june2017/emerging.pdf

Godwin-Jones, R. (2018). Chasing the butterfly effect: Informal language learning online as a complex system. *Language Learning & Technology, 22*(2), 8–27. 10125/44643

Godwin-Jones, R. (2019). Telecollaboration as an approach to developing intercultural communication competence. *Language Learning & Technology, 23*(3), 8–28. http://hdl.handle.net/10125/44691

Godwin-Jones, R. (2021). Big data and language learning: Opportunities and challenges. *Language Learning & Technology, 25*(1), 4–19. http://hdl.handle.net/10125/44747

Godwin-Jones, R. (2023a). Emerging spaces for language learning: AI bots, ambient intelligence, and the metaverse. *Language Learning & Technology, 27*(2), 6–27. https://hdl.handle.net/10125/73501

Godwin-Jones, R. (2023b). Presence and agency in real and virtual spaces: The promise of extended reality for language learning. *Language Learning & Technology, 27*(3), 6–26. https://hdl.handle.net/10125/73529

Godwin-Jones, R. (2024a). Augmented reality as a pedagogical technique. In C. Chapelle (Ed.), *Encyclopedia of applied linguistics* (2nd ed., pp. 1–7). Wiley Blackwell. https://doi.org/10.1002/9781405198431.wbeal20481

Godwin-Jones, R. (2024b). Distributed agency in language learning and teaching through generative AI. *Language Learning & Technology, 28*(2), 5–30. 10125/73570

Godwin-Jones, R., O'Neill, E., & Ranalli, J. (2024). Integrating AI tools into instructed second language acquisition. In C. A. Chapelle, G. H. Beckett, & J. Ranalli (Eds.), *Exploring artificial intelligence in applied linguistics* (pp. 9–23). Iowa State University Digital Press. https://doi.org/10.31274/isudp.2024.154.02

Grant, S. (2025). A comparison of voice-based versus text-based communication with non-character players during complex tasks on the 3D virtual world platform of Second Life: Preliminary student feedback and design implications. *AI-mediated language learning in the metaverse era* (pp. XX–XX). Springer.

Grice, P. (1989). *Studies in the way of words*. Harvard University Press.

Guerrettaz, A. M., Engman, M. M., & Matsumoto, Y. (2021). Empirically defining language learning and teaching materials in use through sociomaterial perspectives. *The Modern Language Journal, 105*(S1), 3–20. https://doi.org/10.1111/modl.12691

Harnad, S. (2024). Language writ large: LLMs, ChatGPT, grounding, meaning and understanding. *arXiv preprint*arXiv:2402.02243. https://doi.org/10.48550/arXiv.2402.02243

Hellmich, E., & Vinall, K. (2021). FL instructor beliefs about machine translation: Ecological insights to guide research and practice. *International Journal of Computer-Assisted Language Learning and Teaching (IJCALLT), 11*(4), 1–18. https://doi.org/10.4018/ijcallt.2021100101

Holden, C. L., & Sykes, J. M. (2011). Leveraging mobile games for place-based language learning. In P. Felicia (Ed.), *Developments in current game-based learning design and deployment* (pp. 27–45). IGI Global. https://doi.org/10.4018/978-1-4666-1864-0.ch003

Hsiao, J.-C., Chang, S. & Chang, J.-S. (2025). AI for language learning: Opportunities and challenges. *AI-mediated language learning in the metaverse era* (pp. XX–XX). Springer.

Hubbard, P. (2019). Five keys from the past to the future of CALL. *International Journal of Computer-Assisted Language Learning and Teaching (IJCALLT), 9*(3), 1–13. https://doi.org/10.4018/ijcallt.2019070101

Hulstijn, J. H. (2005). Theoretical and empirical issues in the study of implicit and explicit second-language learning: Introduction. *Studies in Second Language Acquisition, 27*(2), 129–140. https://doi.org/10.1017/s0272263105050084

Huynh-The, T., Pham, Q. V., Pham, X. Q., Nguyen, T. T., Han, Z., & Kim, D. S. (2023). Artificial intelligence for the metaverse: A survey. *Engineering Applications of Artificial Intelligence, 117*, Article 105581. https://doi.org/10.1016/j.engappai.2022.105581

Jacob, S., Tate, T., & Warschauer, M. (2023). Emergent ai-assisted discourse: Case study of a second language writer authoring with ChatGPT. *arXiv preprint*arXiv:2310.10903. https://doi.org/10.48550/arXiv.2310.10903

Jeong, H., & Li, P. (2023). Neurocognition of social learning of second language: How can second language be learned as first language? In *The Routledge handbook of second language acquisition and neurolinguistics* (pp. 217–229). Routledge. https://doi.org/10.4324/9781003190912-20

Jeong, H., Sugiura, M., Sassa, Y., Wakusawa, K., Horie, K., Sato, S., & Kawashima, R. (2010). Learning second language vocabulary: Neural dissociation of situation-based learning and text-based learning. *NeuroImage, 50*(2), 802–809. https://doi.org/10.1016/j.neuroimage.2009.12.038

Karimi, H., Sañosa, D. J., Rios, K. H., Tran, P., Chun, D. M., Wang, R., & Arya, D. J. (2023). Building a city in the sky: Multiliteracies in immersive virtual reality. *CALICO Journal, 40*(1), 24–44. https://doi.org/10.1558/cj.22838

Kelly, K. (2019, February 12). AR will spark the next big tech platform—Call it mirrorworld. *Wired*. https://www.wired.com/story/mirrorworld-ar-next-big-tech-platform/

Kern, R. (2015). *Language, literacy, and technology*. Cambridge University Press. https://doi.org/10.1017/cbo9781139567701

Kern, R. (2018). Five principles of a relational pedagogy: Integrating social, individual, and material dimensions of language use. *Journal of Technology and Chinese Language Teaching, 9*(2), 1–14. http://www.tclt.us/journal/2018v9n2/kern.pdf

Klekovkina, V., & Denié-Higney, L. (2022). Machine translation: Friend or foe in the language classroom? *L2 Journal, 14*(1), 105–135. https://doi.org/10.5070/l214151723

Klimanova, L. (2021). The evolution of identity research in CALL: From scripted chatrooms to engaged construction of the digital self. *Language Learning & Technology, 25*(3), 186–204. http://hdl.handle.net/10125/73455

Klimanova, L., & Lomicka, L. (2023). Semiotics in CALL: Signs, meanings and multimodality in digital spaces. *Language Learning & Technology, 27*(2), 1–5. https://hdl.handle.net/10125/73500

Kosinski, M. (2023). *Theory of mind might have spontaneously emerged in large language models.* arXiv:2302.0208. https://arxiv.org/abs/2302.02083

Kostka, I., & Toncelli, R. (2023). Exploring applications of ChatGPT to English language teaching: Opportunities, challenges, and recommendations. *Tesl-Ej, 27*(3), 1–19. http://tesl-ej.org/pdf/ej107/int.pdf

Kramsch, C. (2023). Global English: The Indispensable Bridge in Intercultural Communication?. *TEANGA, the Journal of the Irish Association for Applied Linguistics, 30*(1), 1-32. https://journal.iraal.ie/index.php/teanga/article/view/6797

Kronenberg, F., & Poole, F. (2022, November 3). From buzzword to the classroom—Exploring VR gaming for language learning. *The FLTMag.* https://fltmag.com/from-buzzword-to-the-classroom-exploring-vr-gaming-for-language-learning/

Kuhl, P. K., Tsao, F. M., & Liu, H. M. (2003). Foreign-language experience in infancy: Effects of short-term exposure and social interaction on phonetic learning. *Proceedings of the National Academy of Sciences, 100*(15), 9096–9101. https://doi.org/10.1073/pnas.1532872100

Lan, Y-J. (2025). Language learning in the metaverse: Research trends, opportunities and challenges. *AI-mediated language learning in the metaverse era* (pp. XX–XX). Springer.

Lan, Y.-J., & Chen, N.-S. (2024). Teachers' agency in the era of LLM and generative AI: Designing pedagogical AI agents. *Educational Technology & Society, 27*(1), I–XVIII. https://doi.org/10.30191/ETS.202401_27(1).PP01

Lantolf, J. P., Thorne, S. L., & Poehner, M. E. (2014). Sociocultural theory and second language development. In B. VanPatten, G. Keating, & S. Wulff (Eds.), *Theories in second language acquisition* (pp. 221–240). Routledge. https://doi.org/10.4324/9780429503986-10

Larsen-Freeman, D. (1997). Chaos/complexity and second language acquisition. *Applied Linguistics, 18*, 141–165. https://doi.org/10.1093/applin/18.2.141

Larsen-Freeman, D. (2019). On language learner agency: A complex dynamic systems theory perspective. *The Modern Language Journal, 103*, 61–79. https://doi.org/10.1111/modl.12536

Larsen–Freeman, D. (2022). Combinations and connections: Reaching across disciplinary boundaries. *The Modern Language Journal, 106*(S1), 132–140. https://doi.org/10.1111/modl.12753

Lee, S. H., & Wang, S. (2023). Do language models know how to be polite? *Proceedings of the Society for Computation in Linguistics, 6*(1), 375–378. https://doi.org/10.7275/8621-5w02

Legault, J., Fang, S. Y., Lan, Y. J., & Li, P. (2019a). Structural brain changes as a function of second language vocabulary training: Effects of learning context. *Brain and Cognition, 134*, 90–102. https://doi.org/10.1016/j.bandc.2018.09.004

Legault, J., Zhao, J., Chi, Y.-A., Chen, W., Klippel, A., & Li, P. (2019b). Immersive virtual reality as an effective tool for second language vocabulary learning. *Languages, 4*(1), 1–32. https://doi.org/10.3390/languages4010013

Lenat, D. B. (1995). CYC: A large-scale investment in knowledge infrastructure. *Communications of the ACM, 38*(11), 33–38. https://doi.org/10.1145/219717.219745

Levine, G. (2020). A human ecological language pedagogy. *Modern Language Journal, 104*(S1), 1–130. https://doi.org/10.1111/modl.12608

Li, P., & Jeong, H. (2020). The social brain of language: Grounding second language learning in social interaction. *NPJ Science of Learning, 5*(1), 1–9. https://doi.org/10.1038/s41539-020-0068-7

Lin, T.-J., & Lan, Y.-J. (2015). Language learning in virtual reality environments: Past, present, and future. *Educational Technology & Society, 18*(4), 486–497. https://www.jstor.org/stable/jeductechsoci.18.4.486

Lloyd, A., Rogerson, S., & Stead, G. (2017). Imagining the potential for using virtual reality technologies in language learning. In M. Carrier, R. M. Damerow, & K. M. Bailey (Eds.), *Digital language learning and teaching: Research, theory, and practice* (pp. 222–234). Routledge. https://doi.org/10.4324/9781315523293-19

Lorentzen, A., & Bonner, E. (2023, February 12). Customizable ChatGPT AI chatbots for conversation practice. *FLTMag*. https://fltmag.com/customizable-chatgpt-ai-chatbots-for-conversation-practice/

Lv, Z. (2023). Generative artificial intelligence in the metaverse era. *Cognitive Robotics, 3*, 208–217. https://doi.org/10.1016/j.cogr.2023.06.001

Ma, Q., & Yan, J. (2022). How to empirically and theoretically incorporate digital technologies into language learning and teaching. *Bilingualism: Language and Cognition, 25*(3), 392–393. https://doi.org/10.1017/s136672892100078x

Malach, E. (2023). Auto-regressive next-token predictors are universal learners. *arXiv preprint*arXiv:2309.06979. https://arxiv.org/pdf/2309.06979.pdf

Mayer, R. E. (2020). *Multimedia learning*. Cambridge University Press. https://doi.org/10.1017/9781316941355

McConachy, T. (2019). L2 pragmatics as 'intercultural pragmatics': Probing sociopragmatic aspects of pragmatic awareness. *Journal of Pragmatics, 151*, 167–176. https://doi.org/10.1016/j.pragma.2019.02.014

Metz, C. (2024, January 8). Robots learn, Chatbots visualize: How 2024 will be A.i.'s 'leap forward'. *New York Times*. https://www.nytimes.com/2024/01/08/technology/ai-robots-chatbots-2024.html

Michelson, K., & Petit, E. (2017). Becoming social actors: Designing a global simulation for situated language and culture learning. In S. Dubreil & S. L. Thorne (Eds.), *Engaging the world: Social pedagogies and language learning* (pp. 138–167). Cengage. https://scholarspace.manoa.hawaii.edu/server/api/core/bitstreams/2230fa1a-98c5-416b-8f18-938ce62d3e65/content

Mills, N., Courtney, M., Dede, C., Dressen, A., & Gant, R. (2020). Culture and vision in virtual reality narratives. *Foreign Language Annals, 53*(4), 733–760. https://doi.org/10.1111/flan.12494

Mohamed, S., Png, M. T., & Isaac, W. (2020). Decolonial AI: Decolonial theory as sociotechnical foresight in artificial intelligence. *Philosophy & Technology, 33*(4), 659–684. https://doi.org/10.1007/s13347-020-00405-8

Mok, A. (2023, November 12). *Getting emotional with ChatGPT could get you the best outputs*. Business Insider. https://www.businessinsider.com/chatgpt-llm-ai-responds-better-emotional-language-prompts-study-finds-2023-11

Moreno, A. I., & Traxler, J. (2016). MALL-based MOOCs for language teachers: Challenges and opportunities. *Porta Linguarum Monograph, 1*, 73–85. https://doi.org/10.30827/digibug.54090

Nicolaidou, I., Pissas, P., & Boglou, D. (2021). Comparing immersive Virtual Reality to mobile applications in foreign language learning in higher education: A quasi-experiment. *Interactive Learning Environments, 31*(4), 2001–2015. https://doi.org/10.1080/10494820.2020.1870504

O'Gieblyn, M. (2023, May 23). Does AI have a subconscious? *Wired*. https://www.wired.com/story/does-ai-have-a-subconscious/

O'Neill, E. M. (2019). Training students to use online translators and dictionaries: The impact on second language writing scores. *International Journal of Research Studies in Language Learning, 8*(2), 47–65. https://doi.org/10.5861/ijrsll.2019.4002

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

Paivio, A. (1990). *Mental representations: A dual coding approach*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195066661.001.0001

Palacio, R. J. (2012). *Wonder.* Alfred A. Knopf.

Park, G. (2020, April 17). Silicon Valley is racing to build the next version of the Internet. Fortnite might get there first. *The Washington Post*. https://www.washingtonpost.com/video-games/2020/04/17/fortnite-metaverse-new-internet/

Pegrum, M. (2021). Augmented reality learning: Education in real-world contexts. In T. Beaven & F. Rosell-Aguilar (Eds.), *Innovative language pedagogy report* (pp. 115–120). Research-publishing.net. https://doi.org/10.14705/rpnet.2021.50.1245

Peña-Acuña, B., & Navarro-Martínez, Ó. (2023). The promotion of originality perceived in two multimodal storytelling applications: Storybird and scratch. *Education Sciences, 14*(1), 21. https://doi.org/10.3390/educsci14010021

Plass, J., & Jones, L. C. (2005). Multimedia learning in second language acquisition. In R. Mayer (Ed.), *Cambridge handbook of multimedia learning* (pp. 467–488). Cambridge University Press. https://doi.org/10.1017/cbo9780511816819.030

Potamianos, A. (2023, September 12). *Emotion AI and the Ubuntu philosophy*. Medium. https://apotam.medium.com/emotion-ai-and-the-ubuntu-philosophy-c21cf382dcad

Qi, G. Y. (2025). Humanising language learning in the emerging metaverse: A case study of telecollaboration. *AI-mediated language learning in the metaverse era* (pp. XX–XX). Springer.

Qiu, Z., Duan, X., & Cai, Z. G. (2023, May 12). Pragmatic implicature processing in ChatGPT. *PsyArXiv Preprints*. https://doi.org/10.31234/osf.io/qtbh9

Qureshi, N. (2023, May 25). Waluigi, Carl Jung, and the case for moral AI. *Wired*. https://www.wired.com/story/waluigi-effect-generative-artificial-intelligence-morality/

Ranalli, J. (2021). L2 student engagement with automated feedback on writing: Potential for learning and issues of trust. *Journal of Second Language Writing, 52*, Article 100816. https://doi.org/10.1016/j.jslw.2021.100816

Reinhardt, J., & Thorne, S. L. (2020). Digital games as language-learning environments. In J. L. Plass, R. E. Mayer, & B. D. Homer (Eds.), *Handbook of game-based learning* (pp. 409–435). MIT Press.

Roose, K. (2023, May 30). A.I. poses 'risk of extinction,' industry leaders warn. *New York Times*. https://www.nytimes.com/2023/05/30/technology/ai-threat-warning.html

Ruis, L., Khan, A., Biderman, S., Hooker, S., Rocktäschel, T., & Grefenstette, E. (2023, December). The goldilocks of pragmatic understanding: Fine-tuning strategy matters for implicature resolution by LLMs. In A. Oh, T. Naumann, A. Globerson, K. Saenko, M. Hardt, & S. Levine (Eds.), *Proceedings of the 37th International Conference on Neural Information Processing Systems* (pp. 20827–20905). ACM. https://doi.org/10.48550/arXiv.2210.14986

Ryan, M., & Kautzman, K. (Eds.). (2022). *Beyond the traditional essay: Increasing student agency in a diverse classroom with nondisposable assignments*. Vernon Press.

Sadri, F. (2011). Ambient intelligence: A survey. *ACM Computing Surveys (CSUR), 43*(4), 1–66. https://doi.org/10.1145/1978802.1978815

Schwienhorst, K. (2002). Why virtual, why environments? Implementing virtual reality concepts in computer-assisted language learning. *Simulation & Gaming, 33*(2), 196–209. https://doi.org/10.1177/1046878102332008

Sockett, G. (2014). *The online informal learning of English*. Springer. https://doi.org/10.1057/9781137414885

Sørensen, T. F. (2013). We have never been Latourian: Archaeological ethics and the posthuman condition. In *Revisiting actor-network theory in education* (pp. 18–35). Routledge. https://doi.org/10.4324/9781315114521-2

Stephenson, N. (1992). *Snow crash*. Bantam Books.

Su, D., & Goslar, K. (2023, October 19–21). *Evaluating pragmatic competence of artificial intelligence with the lens concept: ChatGPT-4 for Chinese L2 teaching* [Conference presentation abstract]. 20th Annual Technology for Second Language Learning Conference, Ames, IA, United States. https://apling.engl.iastate.edu/wp-content/uploads/sites/221/2023/10/TSLL-2023-Program-Book.pdf

Sykes, J. M., Oskoz, A., & Thorne, S. L. (2008). Web 2.0, synthetic immersive environments, and mobile resources for language education. *CALICO Journal, 25*(3), 528–546. https://doi.org/10.1558/cj.v25i3.528-546

Taguchi, N. (2021). Learning and teaching pragmatics in the globalized world: Introduction to the special issue. *The Modern Language Journal, 105*(3), 615–622. https://doi.org/10.1111/modl.12716

The New London Group. (1996). A pedagogy of multiliteracies: Designing social factors. *Harvard Educational Review, 66*(1), 60–93. https://doi.org/10.17763/haer.66.1.17370n67v22j160u

Thorne, S. L. (2013). Language learning, ecological validity, and innovation under conditions of superdiversity. *Bellaterra Journal of Teaching & Learning Language & Literature, 6*(2), 1–27. https://doi.org/10.5565/rev/jtl3.526

Thorne, S. L., Hellermann, J., & Jakonen, T. (2021). Rewilding language education: Emergent assemblages and entangled actions. *The Modern Language Journal, 105*(S1), 106–125. https://doi.org/10.1111/modl.12687

Tseng, W., & Warschauer, M. (2023). AI-writing tools in education: If you can't beat them, join them. *Journal of China Computer-Assisted Language Learning, 3*(2), 258–262. https://doi.org/10.1515/jccall-2023-0008

Tu, Y.-F., Liu, C.-C., & Chen, J. (2025). Framework and potential applications of Gai-based language learning design. *AI-mediated language learning in the metaverse era* (pp. XX–XX). Springer.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, Ł., & Polosukhin, I.. (2017). Attention is all you need. *Advances in neural information processing systems, 30*. https://proceedings.neurips.cc/paper/2017/file/3f5ee243547dee91fbd053c1c4a845aa-Paper.pdf

Vázquez, C., Xia, L., Aikawa, T., & Maes, P. (2018, July). Words in motion: Kinesthetic language learning in virtual reality. In R. Huang, M. Chang, Kinshuk, N.-S. Chen, D. G. Sampson, S. Murthy, & K. Moudgalya (Eds.), *2018 IEEE 18th International Conference on advanced learning technologies (ICALT)* (pp. 272–276). IEEE. https://doi.org/10.1109/icalt.2018.00069

Vinall, K., & Hellmich, E. A. (2022). Do you speak translate? Reflections on the nature and role of translation. *L2 Journal, 14*(1), 4–25. https://doi.org/10.5070/l214156150

Warschauer, M. (2004). *Technology and social inclusion: Rethinking the digital divide*. MIT Press. https://doi.org/10.7551/mitpress/6699.001.0001

Warschauer, M., Tseng, W., Yim, S., Webster, T., Jacob, S., Du, Q., & Tate, T. (2023). The affordances and contradictions of AI-generated text for second language writers. *Journal of Second Language Writing, 62*. https://doi.org/10.1016/j.jslw.2023.101071

Wei, J., Wang, X., Schuurmans, D., Bosma, M., Xia, F., Chi, E., Le, Q. V., & Zhou, D. (2022). Chain-of-thought prompting elicits reasoning in large language models. *Advances in Neural Information Processing Systems,35*, 24824–24837. https://proceedings.neurips.cc/paper_files/paper/2022/file/9d5609613524ecf4f15af0f7b31abca4-Paper-Conference.pdf

Woo, D. J., Susanto, H., Yeung, C. H., Guo, K., & Fung, A. K. Y. (2023). Exploring AI-generated text in student writing: How does AI help? *arXiv preprint*arXiv:2304.02478. https://arxiv.org/pdf/2304.02478.pdf

Wu, J., & Chen, D. T. V. (2020). A systematic review of educational digital storytelling. *Computers & Education, 147*, Article 103786. https://doi.org/10.1016/j.compedu.2019.103786

Xu, Q., Peng, Y., Wu, M., Xiao, F., Chodorow, M., & Li, P. (2023). *Does conceptual representation require embodiment? Insights from large language models.* arXiv:2305.19103 https://arxiv.org/abs/2305.19103

Yeh, H.-C., Tseng, S.-S., & Heng, L. (2022). Enhancing efl students' intracultural learning through virtual reality. *Interactive Learning Environments, 30*(9), 1609–1618. https://doi.org/10.1080/10494820.2020.1734625

Yeh, Y. L., & Lan, Y. J. (2018). Fostering student autonomy in English learning through creations in a 3D virtual world. *Educational Technology Research and Development, 66*, 693–708. https://doi.org/10.1007/s11423-017-9566-6

Yu, C., & Smith, L. B. (2016). The social origins of sustained attention in one-year-old human infants. *Current Biology, 26*(9), 1235–1240. https://doi.org/10.1016/j.cub.2016.03.026

Robert Godwin-Jones is Professor in the School of World Studies at Virginia Commonwealth University (VCU), where he teaches courses in German and International Studies. He received his Ph.D. in Comparative Literature from the University of Illinois. At VCU, he has served as Chair of the Department of Foreign Languages, Director of the Instructional Development Center (Office of Information Technology), and Director of the English Language Program (Office of International Education). He has served as a guest lecturer at universities in China, France, Germany, and India. His research is principally in applied linguistics, in the areas of language learning and technology and intercultural communication. He writes a regular column for the journal *Language Learning & Technology* on emerging technologies. Robert has published five books and numerous articles and book chapters, as well as regularly presenting at international conferences.