# AI and VR converge: The future of language learning in an emerging metaverse

Robert Godwin-Jones, Virginia Commonwealth University

**Abstract**

The idea of a metaverse holds considerable promise for language learning. The phenomenon of an integrated digital and real 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.

## 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, in particular 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, 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.

**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*, 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 *ChatGPT 5* (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; Kostka & Toncelli, 2023; Lorentzen & Bonner, 2023). The major AI tools (*ChatGPT*, *Bard*, *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, 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 a needed 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, aids 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, Fang et al., 2019; 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 (2024; 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. (2024; 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 towards 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" web site 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., 2024; 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. (2024; this volume).

**AI's language abilities**

It has been demonstrated that 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 were 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* is bland and uninspired compared to that of human writers (Elam, 2023; Kramsch, 2024; 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 (2024, 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 and Goslar, 2023). *ChatGPT* has been shown to be weak in interpreting implied messaging (Ruis et al., 2024; Qiu et al., 2024). 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 Garcia-Pastor, 2020 and McConachy, 2019). The social use of language is dynamic, contingent, and personal. It can also be unpredictable and variable, so that in fact there are 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 wellequipped 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 that 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 human “reinforcement training” is used to fine-tune the system and seek to eliminate errors. The inside workings of AI remain largely a black box (see O’Giebyn, 2023). Becoming aware of that fact is fundamental to developing an understanding of generative AI and therefore a step towards 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.

**Towards 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 real 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 *life-logging* (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 is 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). 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, 2024, 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, the mix of implicit and explicit learning can be achieved, a combination seen 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 towards 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 smartglasses, such as those announced by Google in 2022 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 having 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, this volume; Jeong & Li, 2023).

**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 are 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, 2024, 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., 2019). 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., 2024; 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., 2024, 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., 2024; 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.

**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., 2024, 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, 2024, 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 mostly 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, 2024, 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 *co-presence*) 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 homogenous 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 nonhuman 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, 2018). Learners should be optimally treated, as "developing selves" (Qi 2024, 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 AI 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, 2024; this volume). Interestingly, virtual exchange recently has seen promising examples of VR integration (Baralt et al., 2022; Yeh et al., 2022).

**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 is 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 &

Lan, 2022). The personalization possible through AI, along with the attractive multimodality of VR, are 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, 2024, this volume). This is the future of intelligent, contextualized learning (Lan, 2024, 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 humancentric activities (see Aslam & Li, 2024; 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 nonhuman elements in nature and the environment (Qi, 2024, this volume). Interestingly, postcolonial studies have begun to look at AI from such a perspective (Adams, 2021; Mohammed 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 nonhuman, 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 nonhumans 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., 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 nonhuman actors as important contributors to communication but stresses the social essentialness of human speech. That aligns with the Māori concept of *Whakawhanaungatang*, as discussed in Qi (2024, 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**


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. and Herrera, F. (2023). Explainable Artificial Intelligence (XAI): What we know and what is left to attain Trustworthy Artificial Intelligence. *Information fusion*, *99*, 101805. https://doi.org/10.1016/j.inffus.2023.101805

Aslam, M. S. & Li, Y.-H. (2024). Revolutionizing Language Learning: The Convergence of AI and the Metaverse in Educational Paradigms. *AI-Mediated Language Learning in the Metaverse Era* (pp. XX-XX). New York: 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

Barattieri di San Pietro, C., Frau, F., Mangiaterra, V., Bambini, V., The pragmatic profile of ChatGPT: Assessing the communicative skills of a conversational agent, Sistemi Intelligenti, 2023, 35(2), 379-399, doi: 10.1422/108136

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. http://dx.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 Laguage 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.-J., Hsieh, F.-C., & Yen, Y.-C. (2024). Applying Artificial Intelligence Technologies in Second Language Learning. *AI-Mediated Language Learning in the Metaverse Era* (pp. XX-XX). New York: Springer.

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, N.-S. & Yeh, H.-C. (2024). 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). New York: 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
Chun, D., Kaplan-Rakowski, R., Meyr, J., Ovsiannikova, U., Thrasher, T., & Yuan, Y. (2024). AI-mediated High-Immersion Virtual Reality for Language Learning. *AI-Mediated Language Learning in the Metaverse Era* (pp. XX-XX). New York: 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
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
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, R. (2009). *Implicit and explicit knowledge in second language learning, testing and teaching* (Vol. 42). Bristol, UK: Multilingual Matters. https://doi.org/10.21832/9781847691767-003
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
García-Pastor, M. D. (2020). Researching identity and L2 pragmatics in digital stories. *CALICO Journal*, *37*(1), 46-65. https://www.jstor.org/stable/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. http://dx.doi.org/10125/44268
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. https://doi.org/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. Chapelle, C. (Ed.), *Encyclopedia of Applied Linguistics 2nd Edition* (pp. 1-7). New York: Wiley Blackwell. http://dx.doi.org/10.1002/9781405198431.wbeal20481
Godwin-Jones, R. (2024b). Distributed agency in language learning and teaching through generative AI. *Language Learning & Technology*, ***XX*****(X), XX–XX**. https://doi.org/10125/XXXXX
Grant, S. (2024). 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). New York: 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*.
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). Hershey, Pennsylvania: IGI Global. https://doi.org/10.4018/978-1-4666-1864-0.ch003
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
Hsiao, J.-C., Chang, S. & Chang, J.-S. (2024). AI for language learning: Opportunities and challenges. *AI-Mediated Language Learning in the Metaverse Era* (pp. XX-XX). New York: Springer.
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*, 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). London: 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
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., & 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
Lan, Y-J. (2024). Language Learning in the Metaverse: Research Trends, Opportunities and Challenges. *AI-Mediated Language Learning in the Metaverse Era* (pp. XX-XX). New York: Springer.
Lantolf, J. P., Thorne, S. L., & Poehner, M. E. (2014). Sociocultural theory and second language development. In *Theories in second language acquisition* (pp. 221-240). London: 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. (2019). 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. (2019). 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). London: 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, UK: 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 Tines*. 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). Boston: 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
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
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
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*, 1–15. Advance online publication. 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
Qi, G. (2024). 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. https://doi.org/10.31234/osf.io/qtbh9
Palacio, R. J. (2012). *Wonder.* Alfred A. Knopf.

Paivio, A. (1990). *Mental representations: A dual coding approach*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195066661.001.0001
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 *Cambridge Handbook of Multimedia Learning* (pp. 467-488). Cambridge, UK: 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
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*, 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). Cambridge, Massachusetts: MIT Press.
Roose, K (2023, May 30). A.I. Poses ‘Risk of Extinction,’ Industry Leaders Warn. New York Tines. https://www.nytimes.com/2023/05/30/technology/ai-threat-warning.html
Ruis, L., Khan, A., Biderman, S., Hooker, S., Rocktäschel, T., & Grefenstette, E. (2024). The Goldilocks of Pragmatic Understanding: Fine-Tuning Strategy Matters for Implicature Resolution by LLMs. *Advances in Neural Information Processing Systems*, *36*.
Ryan, M., & Kautzman, K. (Eds.). (2022). *Beyond the Traditional Essay: Increasing Student Agency in a Diverse Classroom with Nondisposable Assignments*. Wilmington, Delaware: 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*. New York: 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). London: Routledge. https://doi.org/10.4324/9781315114521-2
Stephenson, N. (1992). *Snow Crash*. New York: 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

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. (2024). Framework and Potential Applications of Gai-Based Language Learning Design. *AI-Mediated Language Learning in the Metaverse Era* (pp. XX-XX). New York: Springer.

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

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, Ł. and 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

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*. Cambridge, Massachusetts: 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*, 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. http://dx.doi.org/10.1016/j.cub.2016.03.026