Bilingualism: Language and Cognition

cambridge.org/bil

Virtual reality (VR) for language processing and learning: Methods and approaches

Zhe Gao[1], Shaoyun Yu[2], Jiayan Zhao[3], Shufan Yu[1], Suiping Wang[4] and Ping Li[5]

[1]Department of Language Science and Technology, The Hong Kong Polytechnic University, Hong Kong; [2]Department of Linguistics and Modern Languages,The Chinese University of Hong Kong, Hong Kong SAR, China; [3]Laboratory of Geo-Information Science and Remote Sensing,Wageningen University & Research, The Netherlands; [4]Philosophy and Social Science Laboratory of Reading and Development in Children and Adolescents, South China Normal University, China and [5]Division of Social Science, The Hong Kong University of Science and Technology, Hong Kong

Review Article

Cite this article: Gao, Z., Yu, S., Zhao, J., Yu, S., Wang, S. and Li, P. (2026). Virtual reality (VR) for language processing and learning: Methods and approaches. *Bilingualism: Language and Cognition* 1–9. https://doi.org/10.1017/S1366728926101710

Received: 29 November 2025 UTC
Revised: 20 May 2026 UTC
Accepted: 15 June 2026 UTC

Keywords:
virtual reality; multimodal data analytics; embodied cognition; social L2 learning; second language acquisition; neurocognition of language

Corresponding author:
Ping Li;
Email: pingliandrew@ust.hk

Abstract

Virtual reality (VR) is a transformative technology for studying language processing and learning. It offers realistic and naturalistic environments while maintaining experimental control, supporting the exploration of social and sensorimotor cues that have often been overlooked in traditional psycholinguistic research. It can also be integrated with behavioral and neurocognitive measures to examine how learning processes influence learning outcomes. These VR characteristics have brought new opportunities and contributed to the expansion of theoretical frameworks on language processing and learning. Through reviewing recent empirical findings, we establish VR as a novel methodology for investigating how embodied and social experiences shape language processing and learning. We emphasize that realizing this potential fully requires real-time multimodal tracking of learning processes, systematic examination of individual differences and development of richer virtual environments and VR-suitable assessment tasks. This review provides methodological insights for future VR research on the neurocognitive bases of language processing and learning.

Highlights

- We overview how existing studies used virtual reality as an experimental tool to study language processing and learning.
- We explore possible theoretical questions that can be better answered with virtual-reality-based approaches.
- We identify gaps and challenges associated with implementing virtual reality in experimental research.
- We explore possible integrations of virtual reality with behavioral and neurocognitive measures.
- Multimodal approaches provide significant insights into the dynamic language learning processes and outcomes.

1. Introduction

Advances in virtual reality (VR) technology, driven by popular commercial products like Meta Quest and HTC VIVE and the rise of concepts such as the metaverse and spatial computing, have led to its growing popularity in everyday life. This trend presents new opportunities to examine the neurocognitive processes underlying second language (L2) learning and their consequences using VR jointly with behavioral and neurocognitive measures.

VR as a complex media system simulates real-world experiences. It can be delivered via desktop screens, head-mounted displays (HMDs) and cave automatic virtual environment (CAVE) systems. Among them, HMD and CAVE systems provide a full 360° field of regard and enable body-based interactive exploration via real-time motion tracking, resulting in highly immersive experiences. In recent years, researchers have used both HMDs and CAVE in language research, in addition to desktop 3D simulations such as Second Life (Linden Research, Inc.). In this article, unless specified as VR presented on desktops, we focus on the VR technology mediated through HMD and CAVE systems.

VR technology has at least two advantages as an experimental tool. First, the immersive and interactive activities in VR simulate naturalistic environments, which activate learners' sense of presence. This makes the language environments close to daily communication compared to lab-based experiments conducted on 2D screens (Dijkstra & Peeters, 2023), allowing language learners to contextualize the to-be-learned knowledge and retain it in an embodied manner (Casasanto & Jasmin, 2018; Li & Lan, 2022; Pausch et al., 1997). Embodied cognition posits that multimodal information from the environment is an integral part of our linguistic and conceptual representation, and that human cognition develops when the human mind, human body and environments

© The Author(s), 2026. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution-NonCommercial licence (http://creativecommons.org/licenses/by-nc/4.0), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original article is properly cited. The written permission of Cambridge University Press or the rights holder(s) must be obtained prior to any commercial use.

are integrated (Barsalou, 2003, 2008; Glenberg, 1997). VR technology supports this perspective by engaging sensorimotor and social cues, and thus, forms the foundation for developing VR-based experiment platforms. With a proper design, learners can experience processes that are abstract, dynamic, invisible, non-obtainable and/or imaginary to human beings. For example, learners can visualize electrostatics activities (Greenwald et al., 2019) and explore the lunar surface (Paige et al., 2024) in virtual platforms. Such dynamic activities remain physically intangible within traditional classroom settings.

Second, VR provides controlled experimental manipulations while supporting the simultaneous recording of behavior and neurocognitive signals, including eye movements, body movements, electroencephalogram (EEG), functional near infrared spectroscopy (fNIRS) and functional magnetic resonance imaging (fMRI) (Dijkstra & Peeters, 2023). This integration enables multimodal data collection to examine the cognitive and neural bases for both learning processes and outcomes comprehensively. For example, Huizeling et al. (2023) linked eye gazes of participants with EEG data when they talked with virtual interlocutors.

Together, this naturalistic VR-based approach is a game-changer for language research. It allows us to revisit psycholinguistic and neurolinguistic theories on language processing and learning, and test hypotheses that can be evaluated most effectively via this approach (Li & Lan, 2022; Yu et al., 2026). Moreover, it can be used to examine learner characteristics such as cognitive abilities, linguistic backgrounds and language proficiency through not only surveys and tests, but also brain signals such as neural oscillations (Gu et al., 2026; Kepinska et al., 2017).

This article reviews the latest empirical studies on L2 processing and learning of various linguistic constructs, focusing on how VR technology in combination with behavioral and neurocognitive measures can make valuable contributions. Because L2 processing (real-time language comprehension and production) and L2 learning (learning processes over time) address distinct aspects of L2 development, the two topics are reviewed in Section 2 in separate sub-sections, followed by a summary of gaps for future research. The sections on language processing (Section 2.1) and learning (Section 2.2) are structured around three themes: the theoretical frameworks underlying existing work, the ways in which they have used VR as a research tool and the focal factors they have investigated. Section 3 is an overview of the techniques and potential applications of this new approach. The article concludes with a discussion on future directions of VR-based language processing and learning research, as well as digital language learning in general and their contributions to language learning theories (e.g., embodied cognition and social L2 learning). Figure 1 presents an illustration of the integration of VR with behavioral and neurocognitive measures for language processing and learning research.

2. Virtual reality for second language research

2.1. Virtual reality for language processing

Scholars have tested language processing theories in VR, simulating daily oral communication. In a real-life conversation, people rarely hear a sentence with no visual cues from interlocutors and/or

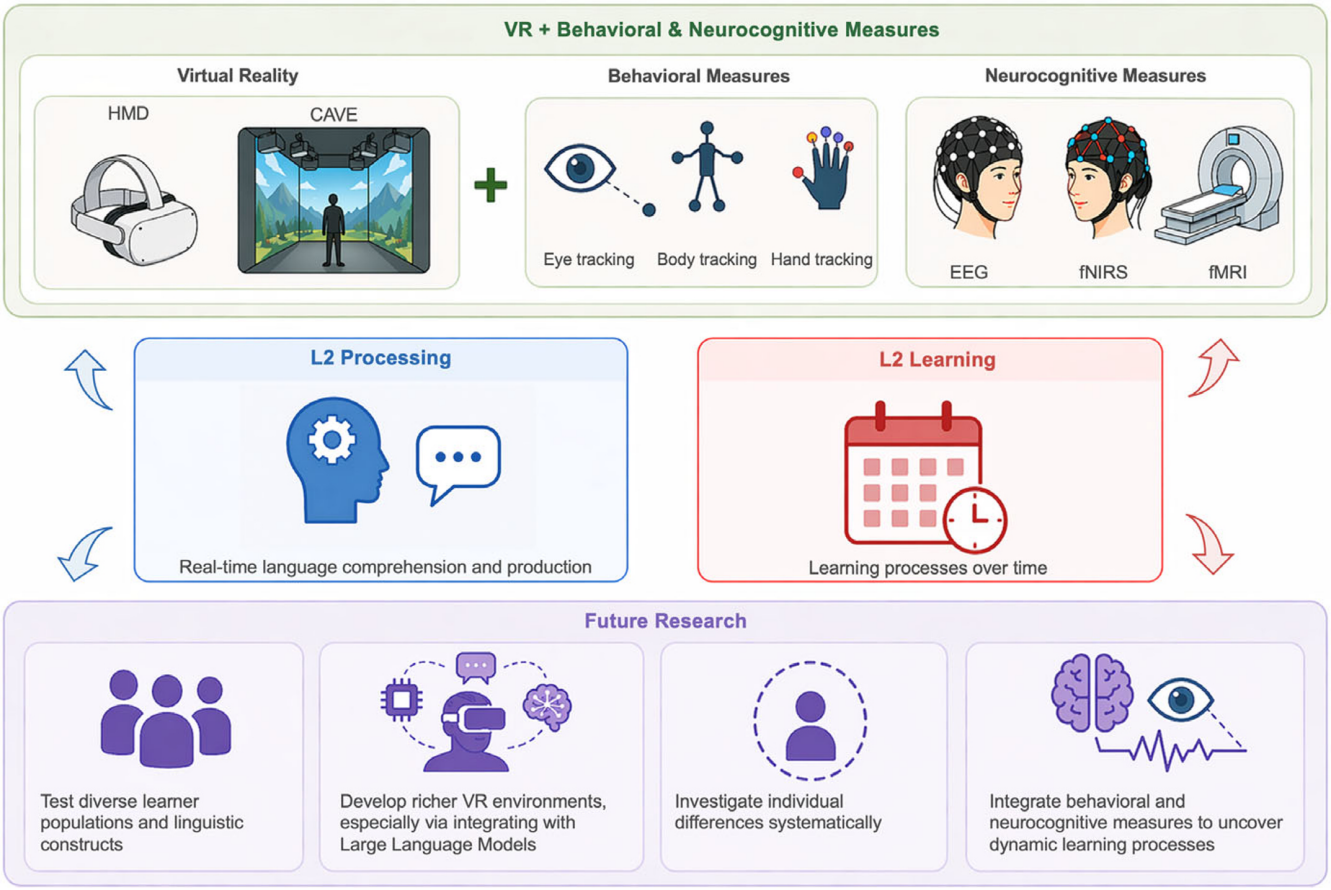


Figure 1. Virtual reality as a novel experimental tool integrated with behavioral and neurocognitive measures (Credit note: the diagram was generated by GPT-5.3 model, and then was edited by the authors for accuracy and clarity).

https://doi.org/10.1017/S1366728926101710 Published online by Cambridge University Press

surrounding environments. VR allows researchers to simulate these conditions while maintaining strict experimental control (Dijkstra & Peeters, 2023; Huizeling et al., 2022; Tromp et al., 2018). Given that this field is still expanding, our discussion includes studies on both native (L1) and L2 processing, focusing on how experiments were conducted with VR and their implications for theories of language processing.

The studies using VR-based paradigms have examined if theories of L2 processing still hold in naturalistic environments. For example, Titus and Peeters (2024) investigated how L2-English speakers with L1-Dutch code-switched between the two languages in a naturalistic setting, i.e., a virtual market in a CAVE system. While the authors found that the overall theoretical claims on inhibitory control still hold (Green & Abutalebi, 2013; Meuter & Allport, 1999), they revealed that, only in the naturalistic virtual environment, switching from L1-Dutch to L2-English was harder than that from L2-English to L1-Dutch. The authors attributed this asymmetry to the naturalistic VR experience, which required more fluent and error-free speech production, when compared with lab-based experiments conducted on 2D screens. As a result, participants needed to exert increased inhibition on their L1 due to their lower proficiency in L2.

In addition, this line of research has examined the effects of non-linguistic sensorimotor information in L1 processing through involving virtual interlocutors to elicit naturalistic responses from the participants (e.g., Huizeling et al., 2022; Misersky et al., 2022; Tromp et al., 2018). For example, using the eye-tracking technique, Huizeling et al. (2022) investigated how native Dutch speakers processed disfluent spoken sentences produced by avatars simulating native Dutch speakers in the CAVE system. As an indicator for anticipating speech in face-to-face conversations in the real world, listeners tend to look at their interlocutor's face when disfluencies occur. The study found that the same strategy was used in the virtual environment, where participants looked at the avatar's face upon hearing the disfluencies while waiting for the forthcoming speech. Likewise, Tromp et al. (2018) employed EEG and observed consistent N400 effects in native Dutch speakers when they listened to sentences that inaccurately described the food on a table in a virtual restaurant. However, it is unclear if L2 learners adopt similar strategies when interacting with avatars in virtual environments.

Although VR as an experimental tool is promising for studying language processing, the linguistic constructs examined with this paradigm are still very limited. There are a few on speech (Huizeling et al., 2022; Titus & Peeters, 2024) and words (Huizeling et al., 2023), but very few studies have investigated the native processing of syntax in virtual environments (Misersky et al., 2022), and even fewer on L2 processing. To go beyond syntax, Taguchi (2022a, 2022b) used VR headsets to examine pragmatics including how L1 and L2 speakers of English used intonation in dispreferred speech acts, such as complaining, in tasks presented via a VR headset or on a 2D screen. In addition, individual characteristics have been barely examined in the existing studies. The combination of the behavioral and neurocognitive measures has also been limited to eye-tracking and EEG.

To summarize, most research in this group suggests that VR can effectively elicit sensorimotor and social information of languages across participant groups even for the information that may be difficult to elicit and/or measure in conventional lab-based experiments. To reach this goal, participants in these studies usually interacted with 3D avatars in virtual environments. Broadening the field of language processing and learning, these studies demonstrate the potential of VR for examining how L2 learners process the multisensory and social information inherent in languages, contexts, and environments. The VR-based paradigms also show potential as pre- and post-test tasks for language learning studies.

2.2. Virtual reality for language learning

While studies on language processing reveal the real-time snapshots of an acquired language, language learning is a developmental process over time. One essential goal for studies on L2 learning is to understand the cognitive and neural bases of this process. Embodied cognition theories have proposed that human cognition, including L2 learning, develops when the body, mind and environments interact with each other and integrate information for learning (Atkinson, 2010; Barsalou, 2003, 2008; Glenberg, 1997). Under this framework, social L2 learning theory posits that social cues such as joint attention embedded in social interaction promote learning (Li & Jeong, 2020; Zappa et al., 2025). However, it is challenging for traditional lab-based experiments on 2D screens to capture the dynamic interactions. With the help of VR, we can investigate to what extent L2 learners can encode the sensorimotor and social information of language into their mental representations during learning, and to what extent the encoded information enhances language learning.

Existing studies have primarily examined L2 vocabulary learning, with cognitive and neural measures conducted in pre- and post-tests. One line of the studies incorporated EEG during the (pre- and) post-tests to assess learning outcomes and the quality of mental representations of the learned words. In a study by Zappa et al. (2024), novice learners of Serbian with L1-French learned Serbian action verbs through either imitating the actions or pointing to the target verbs on a VR platform. Indicated by the increased N400 effect for semantic processing from the pre-test to the post-test, the authors found that both action methods in VR facilitated semantic encoding of the newly learned Serbian verbs. Further evidence was found by Jiao et al. (2024, 2026) examining if visuospatial cues in virtual environments facilitated the learning of German nouns by L1-Chinese speakers. Their semantic priming task (2025) showed that by eliciting increased N400, VR enhanced the encoding of the learned German nouns in semantic memory.

Another line of research adopted magnetic resonance imaging (MRI) in the pre- and post-tests. Using functional MRI, scholars have shown that vocabulary learned with pictures and gestures, rather than word-word association, elicits greater activation in the visual and motor cortices during retrieval, suggesting more enriched encoding of words (Bechtold et al., 2023; Mayer et al., 2015). Moreover, using structural MRI, Legault et al. (2019b) investigated structural changes in learners' brains under two learning conditions: Second Life, a desktop VR platform and a conventional picture-word association task. Novice learners of Chinese with L1-English learned 90 Chinese words over seven sessions within 20 days in one of the two learning conditions. Learning outcomes were measured at three time points, (1) after each learning session, (2) 3 weeks later and (3) 6 weeks later. MRI scans were obtained after the first and last learning sessions. The authors found that in the VR-based learning condition, the learners' performance was positively correlated with the cortical thickness in the right inferior parietal lobe, a region associated with sensorimotor integration and embodied cognition (Claeys et al., 2003; Macaluso et al., 2003). In contrast, the picture-word association condition was associated with cortical thickness in the right inferior frontal gyrus, indicative of semantic learning (Jones et al., 2012). The findings

https://doi.org/10.1017/S1366728926101710 Published online by Cambridge University Press

suggest that VR-based learning may lead to distinct brain structure changes compared to the picture-word association task.

Different from the research on language processing, recent studies have explored how individual characteristics influence L2 learning with VR-based platforms. Jiao et al. (2024) found that L2 age of acquisition was negatively correlated with N200 amplitude, an indicator of early lexical selection and phonological processing, whereas inhibitory control showed a null correlation. Rudling et al. (2026) revealed that VR-based learning platforms might benefit child L2 learners who struggled with tasks requiring sustained attention. Both Legault et al. (2019a) and Macedonia et al. (2023) reported that less successful learners especially benefited from VR learning. Macedonia et al. (2023) suggest that the impact of VR-based motor cues, such as grasping, depends on a learner's language aptitude, i.e., the specific cognitive abilities for L2 learning, which vary significantly across individual learners (Dörnyei & Skehan, 2003; Rogers et al., 2017). For high-aptitude learners with efficient declarative memory, these procedural cues offer negligible benefits. Conversely, for low-aptitude learners, interactive procedural activities provide a vital compensatory pathway that supplements a less efficient declarative memory system. Although the study did not directly assess the learners' declarative and procedural memory, it represents a crucial step toward understanding how these memory systems as individual characteristics influence L2 learning in virtual environments (Morgan-Short et al., 2014; Saito et al., 2022; Ullman & Lovelett, 2018).

To summarize, sensorimotor and social cues from immersive and interactive environments may facilitate L2 learning and thus contribute to the theoretical frameworks of embodied cognition and social L2 learning (see Mathias & von Kriegstein, 2023; Yu et al., 2026 for detailed discussion). However, the conclusion is tentative given the following limitations. First, the behavioral and neurocognitive measures have been dominantly used in the post-learning testing sessions, rather than both the learning and the testing sessions (except for Gao et al., under review.a; Hsiao et al., 2017; see Section 3.2). Consequently, the learning processes remain a black box. Second, the neurocognitive measures in the testing sessions were based on conventional, non-embodied 2D paradigms, instead of directly assessing how well the sensorimotor and social cues of the learned language are encoded and retained (Jiao et al., 2024, 2026; Lin et al., 2026). It is necessary to develop novel VR-suitable tasks for the pre- and post-tests to accurately test the learning outcomes and to conduct the whole experiment in VR. Third, it is still unexplored how sensorimotor and social cues can maximize learning outcomes, especially in the long term within the constraints of human cognition (Zou et al., 2025). Commercial VR learning platforms as well as other digital language learning tools often include attractive but complex interactive features, but such features are often not tested or validated empirically and hence may be unnecessary or ineffective (Li & Lan, 2022). Future research should identify both effective and distracting design features for learners and integrate evidence-based learning strategies into immersive experiences. Such insights can greatly improve the pedagogical value of VR-based language learning platforms in the long run (Fuhrman et al., 2021; Gao et al., under review.a; Hartfill et al., 2020).

2.3. Gaps and areas for future development

With the above analyses, we have identified four shared areas for future advancement. First, future research should include more diverse learner populations, especially speakers of non-European languages, examine a broader range of linguistic constructs, test more environmental and social contexts and examine the interactions with interlocutors. This will contribute to the discussion on how the sensorimotor and social cues embedded in human languages and virtual environments are processed and learned within the constraints of human cognition and experience, and thus broaden the theories on language processing and learning in naturalistic embodied environments (Barsalou, 2003, 2008; Hagoort, 2023; Mahon & Caramazza, 2008; Wilson, 2002; Zou et al., 2025). For example, we are currently testing if motion and spatial cues facilitate L2 learning of spatial language in a VR environment.

Second, we need to develop richer dynamic virtual environments based upon recent developments in VR technology and its integration with other technologies including Large Language Models (LLMs), portable neuroimaging devices and biofeedback (Yu et al., 2026). While LLMs enable sophisticated, naturalistic human-machine interaction, VR provides an embodied 3D context that grounds these interactions in vivid scenarios. For example, it is now feasible to generate context-appropriate, interlocutor-specific conversations by including LLM-powered AI agents in VR. This will help create a new mode of communication, moving from one-sentence responses to naturalistic human-avatar dialogues (Wan et al., 2024). Additionally, by integrating measures like eye movements (Bovo et al., 2025) and fNIRS (Wen et al., 2025), AI agents and VR environments can be dynamically adjusted in accordance with human real-time responses.

Third, the impacts of individual characteristics warrant comprehensive investigation. On the one hand, we need to examine how the well-established individual cognitive abilities like language aptitude, working memory span (Gao et al., under review.a; Legault et al., 2019a), attention span (Rudling et al., 2026), declarative and procedural memory systems (Ullman & Lovelett, 2018) and linguistic experiences such as length of residence in an L2 environment (Gao et al., under review.b; Saito, 2015) influence L2 processing and learning in naturalistic virtual environments. On the other hand, we need to examine the abilities relevant to VR such as spatial abilities, motion sickness and prior VR experience (Li et al., 2020; Zhao et al., 2021). This line of inquiry will identify which learner groups benefit most from immersive environments, enabling developers to tailor VR platforms to diverse needs. Furthermore, such research deepens our understanding of the neurocognitive mechanisms for language processing and learning. For instance, it allows us to investigate whether established structural brain differences between fast and slow learners persist within ecologically valid, naturalistic settings.

Lastly, in terms of methodology, both language processing and learning research have utilized limited behavioral and neurocognitive measures in VR, although it is crucial to understand the cognitive and neural mechanisms for language processing in virtual environments, and to uncover how neural changes occur during learning with VR and how they affect both short- and long-term learning outcomes. An important direction for the future will be approaches for combining behavioral and neurocognitive measures within VR and for analyzing the multimodal data collected from such measures, to which we now turn.

3. Tools and methods for VR-based approach

3.1. Head-mounted display (HMD) and cave automatic virtual environment (CAVE) system for cognitive studies

Both HMD and CAVE are VR systems that offer users a 360° field of regard in virtual environments (Zhao et al., 2019). The CAVE

https://doi.org/10.1017/S1366728926101710 Published online by Cambridge University Press

system is a cube-shaped room with multiple projection surfaces (e.g., walls and floor) capable of displaying computer-generated 3D imagery (Cruz-Neira et al., 1992; DeFanti et al., 2011). Users inside the CAVE system can simultaneously explore virtual content while seeing each other's bodies and responses. Such a shared space facilitates natural communications and collaborative demonstrations but comes at the high cost of specialized infrastructure. In contrast, HMDs provide a highly immersive, portable and cost-effective solution, allowing individual users to experience virtual environments from anywhere and with minimal visual input from the physical surroundings (Slater & Wilbur, 1997). Both CAVE and HMD systems support interactive exploration through real-time motion tracking, although HMDs typically afford more embodied and tangible experiences due to their seamless integration of a virtual body and hand-based interactions (Zhao et al., 2020). When wearing an HMD, users are visually isolated from the real world around them, which enhances immersion but limits social contact with other users (Lantz, 1996). In this regard, recent advances in networked collaborative VR now bring users from different physical locations into the same virtual environment (e.g., *VRChat*). Combined with fully rigged and customizable avatars, these developments are helping address earlier concerns about isolation among HMD users. Figure 1 presents an illustration of the CAVE and HMD systems and their utility for language learning research.

3.2. Tracking learning processes in virtual reality with behavioral and neurocognitive measures

As VR promises to bridge the gap between controlled experiments and real-world language learning, it becomes equally important to capture the diverse cognitive and behavioral processes that unfold along with learners' exploration and interaction in these immersive spaces (Li et al., 2020; Li & Lan, 2022; Peeters, 2019; Zappa & Frenck-Mestre, 2023). Combining VR with behavioral and neurocognitive measures presents an invaluable opportunity to investigate learners' real-time cognitive states, brain activities and behavioral patterns during ongoing learning processes (i.e., eye-brain-behavior correspondence; see Gu et al., 2024), thus introducing a novel way of uncovering how learning processes influence learning outcomes, which remains rarely studied in experiments with a pretest-training-posttest design, as discussed above. Advanced multimodal data analytic methods such as intersubject correlation (ISC) analysis can be used to handle these complex large datasets (Freund et al., 2013; Madsen & Parra, 2022). In this section, we will discuss possibilities for the use of HMD and CAVE systems to be combined with behavioral and neurological data recordings, limitations of each combination, ways in which they may complement one another and methods for analyzing and interpreting these data.

3.2.1. Eye-tracking

Since VR is inherently a visuospatial experience, eye movements serve as a window into learners' attentional dynamics and other cognitive processes (Gu et al., 2024; Shadiev & Li, 2023). Many high-end VR headsets, such as HTC Vive Pro Eye, Varjo XR series and Meta Quest Pro, feature integrated eye-tracking capabilities, with sampling rates typically ranging from 30 to 200 Hz. These systems continuously record raw eye movement data, including gaze coordinates, gaze direction and pupil dilation, while synchronizing their timestamps with events in the VR environment. As for the CAVE system, participants need to wear transparent stereo glasses with the eye tracking function. The studies discussed in previous sections (e.g., Huizeling et al., 2022, 2023) used SMI Eye-Tracking Glasses 2 Wireless with a binocular sampling rate of 60 Hz (Eichert et al., 2018).

Fixation counts and gaze durations are two of the most widely used eye-tracking metrics in VR-based learning research (Bacca-Acosta et al., 2022; Wang et al., 2021). These metrics can provide precise indicators of visual attention and learning patterns based on the fixation frequency and duration on specific objects in VR. For example, Bacca-Acosta et al. (2022) used the Vive Pro Eye to investigate the effectiveness of scaffolding materials in English learning. They found that an immersive VR environment allowing free exploration, compared to a seated condition, induced significantly more fixations on scaffolding materials, which in turn led to better learning outcomes. Recent analytical methods also enable researchers to move beyond these classical eye-tracking metrics. First, network science methods can model eye movement trajectories as scanpath networks, revealing the effects of individual differences and exploration strategies for effective learning (Ma et al., 2023). Second, entropy-based features can be extracted from eye movement data to estimate the diversity of exploration patterns (Krejtz et al., 2014). For example, inspired by studies on human navigation in cities (Heller et al., 2020) and mice in mazes (Freund et al., 2013), Gao et al. (under review.a) calculated global and local roaming entropies to estimate L2 learners' exploration areas in virtual environments. Their analyses demonstrated that distant environmental cues (indicated by global roaming entropy) facilitated L2 learning of Spanish words whereas adjacent ones (indicated by local roaming entropy) hindered the learning. Finally, machine learning models, such as Eye Movement analysis with Hidden Markov Models, have been successfully employed to classify participants' visual attention patterns in domains such as face recognition and driving (Hsiao et al., 2021; Yang et al., 2023). These advanced methodologies offer promising avenues for a deeper understanding of learners' cognitive processes in VR-based language learning.

3.2.2. Body movement tracking

One of VR's strengths is its ability to incorporate natural body movements, such as walking, reaching, controller clicking and object grasping, into the learning process. As learners navigate in a virtual environment, their movement trajectories can be recorded and analyzed (Brunec et al., 2023). For example, Yang et al. (2021) calculated the movement distance of head and hands to estimate exploration areas. In addition to movement trajectories, controller clicks are especially informative regarding how participants interact with learning materials. Gao et al. (under review.a) quantified learners' exploration paths by calculating the ratio between actual and possible controller clicks among to-be-learned Spanish words in a virtual environment (Borsboom et al., 2021; Ma et al., 2023). They found that linear rather than net-like exploration paths facilitated word learning. In an analysis of the order of clicks data from Lan et al.'s (2015) study, Hsiao et al. (2017) revealed that learners who grouped the to-be-learned words according to their functions, such as their phonological and semantic features, learned better. Some VR headsets such as Meta Quest Pro enable even finer-grained tracking of body movements with built-in hand tracking, eliminating the need for controllers or external motion-capture systems. These developments have opened new opportunities for studying complex body-based interactions and naturalistic hand gestures during learning.

https://doi.org/10.1017/S1366728926101710 Published online by Cambridge University Press

3.2.3. Brain imaging

Observable learning behaviors and outcomes are often underpinned by learners' brain processes. Neural evidence tells us how multimodal information in naturalistic learning settings is integrated in the brain and ultimately contributes to knowledge formation. Therefore, integrating VR with brain imaging technologies can uncover the neural mechanisms of language learning, often in real time and not revealed by behavioral measures or tests. In this section, we focus on three brain imaging technologies, EEG, fNIRS and fMRI, which have been primarily used in neurocognitive research. We exclude structural MRI because it is mainly used to assess brain structure changes, which occur over relatively long-time scales (Legault et al., 2019b) although there is recent evidence that such changes can occur rapidly (Rumetshofer et al., 2026).

Currently, simultaneous VR-brain imaging research on language learning remains scant (Hua & Wang, 2023; Peng et al., 2024). Nevertheless, existing studies on language processing have demonstrated the potential of combining these technologies. EEG has been a popular choice in language comprehension studies due to its high temporal resolution and relatively flexible setup (Huizeling et al., 2023; Tromp et al., 2018; Zappa et al., 2019). For example, Zappa et al. (2019) combined a CAVE system with a 64-channel EEG to investigate embodied cognition during the processing of verbs about hand and arm actions. The analysis of event-related spectral perturbation revealed significant motor-related neural oscillatory activity in the *mu* and *beta* bands, contributing empirical evidence to the embodied cognition theory that motor information can be encoded in mental representations. In the context of language learning, a few studies have successfully combined digital games with portable, single-channel EEG systems (e.g., Hsu, 2022; Yu & Tsuei, 2023). While the limited spatial coverage in the brain may constrain the depth of neurocognitive insights, these studies demonstrate EEG's capability to monitor cognitive engagement and attention during interactive game-based learning.

Researchers have also used fNIRS to investigate language learning in naturalistic interactive contexts, as it can capture hemodynamic activities from language-related cortices while allowing moderate motions (Lawrence et al., 2021; Soltanlou et al., 2018). This makes fNIRS suitable for examining language processing and learning in interactive VR environments, although this combination has been predominantly applied in studies of non-linguistic tasks and cognitive skills (Peng et al., 2024 for a review). Additionally, the hyperscanning technique enables capturing the neural dynamics of multiple participants during interaction. Recent studies demonstrate that discussions elicit high frontal inter-brain coherence, linking synchrony with improved L2 word learning outcomes (e.g., Zhang et al., 2025). Although not yet applied in VR, it is possible to apply hyperscanning for role plays and discussions in virtual environments, investigating how immersion and social dynamics drive L2 learning. fNIRS data are typically analyzed by calculating the changes in oxygenated hemoglobin (HbO). Concentration comparison, activation analysis and functional connectivity are mostly used for a block- or event-related experimental design (Yücel et al., 2021). In contrast, ISC (Hasson et al., 2004; Madsen & Parra, 2022) and wavelet coherence (Zhang et al., 2025) are often used to capture inter-brain coherence during interactions, particularly in the hyperscanning and naturalistic paradigms (see also discussion below on IS-RSA).

Selecting an appropriate brain imaging method for VR-based research requires considering system compatibility as well as trade-offs in temporal resolution, spatial resolution and motion tolerance. Both EEG and fNIRS systems allow participants to use wearable caps, and the caps with low-profile sensors are often compatible with VR headsets and CAVE systems. In contrast, VR–fMRI integration requires specialized MRI-safe devices, such as the VisualSystem HD (Nordic NeuroLab, Norway). EEG offers excellent temporal resolution but poor spatial resolution, and it is prone to non-neural electrical artifacts such as motion-induced noises. In contrast, fMRI has high spatial resolution but poor temporal resolution, and can be unnatural when combined with VR given the constraints on participants' body movements. fNIRS offers moderate temporal and spatial resolutions, positioned between EEG and fMRI, and has a higher motion tolerance than the other two methods.

In Sections 3.2.1 and 3.2.2, we discussed how eye and body movement data in VR could be useful. It is possible to incorporate brain imaging techniques with eye and body movement to gain a multifaceted view of the learning process. Huizeling et al. (2023) explored the possibility of linking eye-tracking and EEG with the CAVE system for studying language prediction. They found that the two measurements can be complementary to each other. For example, when participants stopped looking at the virtual interlocutor as shown by the eye-tracking data, EEG data including theta and alpha frequency bands and the N400 component could indicate whether the participants have lost confidence in their prediction or got distracted. Mixed effects models were used to connect eye movements and N400 amplitude. Though not used in VR yet, Gu et al. (2024) investigated how the presence of instructors in online lectures influenced the student learning outcomes by linking the students' eye movements, brain activities, and post-test performances via ISC and inter-subject representational similarity analyses (IS-RSA). ISC measures similarities of brain activities among the students during the video lectures, while IS-RSA correlates the similarities with those in eye movements (Finn et al., 2020; Hasson, 2025; Madsen & Parra, 2022).

In summary, emerging VR technologies offer sophisticated behavioral tracking capabilities and are compatible with advanced brain imaging methods. By capturing rich, multimodal data throughout the learning process, VR studies of language learning have the potential to transcend the limitations of traditional pre- and post-test assessments and questionnaires. Synthesizing these complex data will yield unprecedented insights into the real-time cognitive and neural mechanisms underlying language processing and learning.

4. Conclusion

This review explores VR as a cutting-edge tool for research on L2 processing and learning. VR offers a dynamic, immersive environment to investigate how multimodal cues and sensorimotor experiences shape L2 knowledge representation, encoding and retention. We highlight innovative approaches that integrate VR with behavioral and neurocognitive measures so as to gain insights into how L2 learning processes impact learning outcomes.

Looking ahead, future research on both language processing and learning with VR-based paradigms should include more diverse learner populations, explore a broader range of linguistic constructs and delve into individual learner characteristics including cognitive abilities, linguistic experiences and VR-specific individual differences. Additionally, we should develop richer virtual environments and integrate with other emerging technologies such as LLMs and physiological monitoring (Yu et al., 2026). We should also develop VR-suitable assessment tasks for pre- and post-tests and integrate

https://doi.org/10.1017/S1366728926101710 Published online by Cambridge University Press

more behavioral and neurocognitive measures with VR, so that we can conduct experiments within virtual environments while collecting multimodal data during both learning and testing periods. Finally, advancing sophisticated multimodal data analysis techniques is essential for tracking learning processes in real time. This requires researchers to adopt analytic methods from other fields such as spatial cognition (Brunec et al., 2023; Zhao et al., 2022).

Digital language learning as a rapidly evolving research field not only enriches our understanding of L2 processing and learning but also provides novel methodologies for understanding how other digital learning tools, such as mobile Apps and online platforms, can transform research on personalized language learning for the future (Li & Lan, 2022). Together, these findings will inform the development of technology-enhanced language learning tools that promote effective language education for learners with diverse individual characteristics.

Data Availability Statement. This study is a review article and did not involve the generation, collection, or analysis of original research materials or datasets. All relevant literature sources are appropriately cited in the References section.

Acknowledgments. Writing of this article and the research reported here was supported by grants from the Hong Kong Research Grants Council (Project # 15607623) and the Sin Wai Kin Foundation to Ping Li (where the author was previously affiliated), from the sector plans of the Dutch Ministry of Education, Culture and Science to Jiayan Zhao and from the Key Research and Development Program of Guangdong, China (Project # 2023B0303010004) to Suiping Wang. Zhe Gao, Shaoyun Yu and Shufan Yu have been supported by Postdoctoral Research Fellowships from the Hong Kong Polytechnic University. We thank Brian MacWhinney and the members of the Brain, Language, and Computation Laboratory for their insightful comments and suggestions on the preparation of the article.

Competing interests. The authors declare that they have no known competing interests that could have appeared to influence the work reported in this article.

Use of AI. We have used AI-assisted technologies in creating this article. We made use of ChatGPT to generate Figure 1. Details have been provided in the figure caption.

References

Atkinson, D. (2010). Extended, embodied cognition and second language acquisition. *Applied Linguistics*, **31**(5), 599–622. https://doi.org/10.1093/applin/amq009

Bacca-Acosta, J., Tejada, J., Fabregat, R., Kinshuk, & Guevara, J. (2022). Scaffolding in immersive virtual reality environments for learning English: An eye tracking study. *Educational Technology Research and Development*, **70**(1), 339–362. https://doi.org/10.1007/s11423-021-10068-7

Barsalou, L. W. (2003). Situated simulation in the human conceptual system. *Language and Cognitive Processes*, **18**(5–6), 513–562. https://doi.org/10.1080/01690960344000026

Barsalou, L. W. (2008). Grounded cognition. *Annual Review of Psychology*, **59**(1), 617–645. https://doi.org/10.1146/annurev.psych.59.103006.093639

Bechtold, L., Cosper, S. H., Malyshevskaya, A., Montefinese, M., Morucci, P., Niccolai, V., Repetto, C., Zappa, A., & Shtyrov, Y. (2023). Brain signatures of embodied semantics and language: A consensus paper. *Journal of Cognition*, **6**(1), 61. https://doi.org/10.5334/joc.237

Borsboom, D., Deserno, M. K., Rhemtulla, M., Epskamp, S., Fried, E. I., McNally, R. J., Robinaugh, D. J., Perugini, M., Dalege, J., Costantini, G., Isvoranu, A. M., Wysocki, A. C., van Borkulo, C. D., van Bork, R., & Waldorp, L. J. (2021). Network analysis of multivariate data in psychological science. *Nature Reviews Methods Primers*, **1**(1), 58. https://doi.org/10.1038/s43586-021-00055-w

Bovo, R., Abreu, S., Ahuja, K., Gonzalez, E. J., Cheng, L.-T., & Gonzalez-Franco, M. (2025). EmBARDiment: An embodied AI agent for productivity in XR. In *2025 IEEE conference virtual reality and 3D user interfaces (VR)* (pp. 708–717), Saint-Malo, France. https://doi.org/10.1109/VR59515.2025.00093

Brunec, I. K., Nantais, M. M., Sutton, J. E., Epstein, R. A., & Newcombe, N. S. (2023). Exploration patterns shape cognitive map learning. *Cognition*, **233**, 105360. https://doi.org/10.1016/j.cognition.2022.105360

Casasanto, D., & Jasmin, K. M. (2018). Virtual reality. In A. de Groot & P. Hagoort (Eds.), *Research methods in psycholinguistics and the neurobiology of language* (pp. 174–189). John Wiley & Sons.

Claeys, K. G., Lindsey, D. T., De Schutter, E., & Orban, G. A. (2003). A higher order motion region in human inferior parietal lobule: Evidence from fMRI. *Neuron*, **40**(3), 631–642. https://doi.org/10.1016/S0896-6273(03)00590-7

Cruz-Neira, C., Sandin, D. J., DeFanti, T. A., Kenyon, R. V., & Hart, J. C. (1992). The CAVE: Audio visual experience automatic virtual environment. *Communications of the ACM*, **35**(6), 64–72. https://doi.org/10.1145/129888.129892

DeFanti, T., Acevedo, D., Ainsworth, R., Brown, M., Cutchin, S., Dawe, G., Doerr, K., Johnson, A., Knox, C., Kooima, R., Kuester, F., Leigh, J., Long, L., Otto, P., Petrovic, V., Ponto, K., Prudhomme, A., Rao, R., Renambot, L., … Wickham, G. (2011). The future of the CAVE. *Open Engineering*, **1**(1), 16–37. https://doi.org/10.2478/s13531-010-0002-5

Dijkstra, T., & Peeters, D. (2023). *The new psychology of language: From body to mental model and back*. Routledge. https://doi.org/10.4324/9781003326274

Dörnyei, Z., & Skehan, P. (2003). Individual differences in second language learning. In C. J. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 589–630). Wiley. https://doi.org/10.1002/9780470756492.ch18

Eichert, N., Peeters, D., & Hagoort, P. (2018). Language-driven anticipatory eye movements in virtual reality. *Behavior Research Methods*, **50**(3), 1102–1115. https://doi.org/10.3758/s13428-017-0929-z

Finn, E. S., Glerean, E., Khojandi, A. Y., Nielson, D., Molfese, P. J., Handwerker, D. A., & Bandettini, P. A. (2020). Idiosynchrony: From shared responses to individual differences during naturalistic neuroimaging. *NeuroImage*, **215**, 116828. https://doi.org/10.1016/j.neuroimage.2020.116828

Freund, J., Brandmaier, A. M., Lewejohann, L., Kirste, I., Kritzler, M., Krüger, A., Sachser, N., Lindenberger, U., & Kempermann, G. (2013). Emergence of individuality in genetically identical mice. *Science*, **340**(6133), 756–759. https://doi.org/10.1126/science.1235294

Fuhrman, O., Eckerling, A., Friedmann, N., Tarrasch, R., & Raz, G. (2021). The moving learner: Object manipulation in virtual reality improves vocabulary learning. *Journal of Computer Assisted Learning*, **37**(3), 672–683. https://doi.org/10.1111/jcal.12515

Gao, Z., Jiang, Y., Hu, Y., Chan, S.-D., & Li, P. (under review.a). Virtual memory palaces for language learning. *A multimodal approach*

Gao, Z., Zhou, X., Xia, Y., & Chen, L. (under review.b). Self-assessed L2 proficiency reflects linguistic experience and objective ability.

Glenberg, A. M. (1997). What memory is for. *Behavioral and Brain Sciences*, **20**(1), 1–19. https://doi.org/10.1017/S0140525X97000010

Green, D. W., & Abutalebi, J. (2013). Language control in bilinguals: The adaptive control hypothesis. *Journal of Cognitive Psychology*, **25**(5), 515–530. https://doi.org/10.1080/20445911.2013.796377

Greenwald, S. W., Corning, W., McDowell, G., Maes, P., & Belcher, J. (2019). ElectroVR: An electrostatic playground for collaborative, simulation-based exploratory learning in immersive virtual reality. In *In proceedings of the international conference on computer supported collaborative learning (CSCL '19)* **2**, (pp. 997–1000)

Gu, C., Peng, Y., & Li, P. (2026). Advances in bilingualism as a dynamic process: 30 years of exploration in bilingual mind and brain. *Journal of Neurolinguistics*, **77**, 101288. https://doi.org/10.1016/j.jneuroling.2025.101288

Gu, C., Peng, Y., Nastase, S. A., Mayer, R. E., & Li, P. (2024). Onscreen presence of instructors in video lectures affects learners' neural synchrony and visual attention during multimedia learning. *Proceedings of the National Academy of Sciences of the United States of America*, **121**(12), e2309054121. https://doi.org/10.1073/pnas.2309054121

Hagoort, P. (2023). The language marker hypothesis. *Cognition*, **230**, 105252. https://doi.org/10.1016/j.cognition.2022.105252

Hartfill, J., Gabel, J., Neves-Coelho, D., Vogel, D., Räthel, F., Tiede, S., Ariza, O., & Steinicke, F. (2020). Word saber: An effective and fun VR vocabulary

learning game. In B. Preim, A. Nürnberger, & C. Hansen (Eds.), *MuC '20: Proceedings of the conference on mensch und computer* (pp. 145–154). Association for Computing Machinery. https://doi.org/10.1145/3404983.3405517

Hasson, U. (2025). Uncovering a timescale hierarchy by studying the brain in a natural context. *Journal of Neuroscience*, **45**(12), e2368242025. https://doi.org/10.1523/JNEUROSCI.2368-24.2025

Hasson, U., Nir, Y., Levy, I., Fuhrmann, G., & **Malach, R.** (2004). Intersubject synchronization of cortical activity during natural vision. *Science*, **303**(5664), 1634–1640. https://doi.org/10.1126/science.1089506

Heller, A. S., Shi, T. C., Ezie, C. E. C., Reneau, T. R., Baez, L. M., Gibbons, C. J., & **Hartley, C. A.** (2020). Association between real-world experiential diversity and positive affect relates to hippocampal–striatal functional connectivity. *Nature Neuroscience*, **23**(7), 800–804. https://doi.org/10.1038/s41593-020-0636-4

Hsiao, I. Y. T., Lan, Y.-J., Kao, C.-L., & **Li, P.** (2017). Visualization analytics for second language vocabulary learning in virtual worlds. *Educational Technology & Society*, **20**(2), 16

Hsiao, J. H., Lan, H., Zheng, Y., & **Chan, A. B.** (2021). Eye movement analysis with hidden Markov models (EMHMM) with co-clustering. *Behavior Research Methods*, **53**(6), 2473–2486. https://doi.org/10.3758/s13428-021-01541-5

Hsu, L. (2022). To CALL or not to CALL: Empirical evidence from neuroscience. *Computer Assisted Language Learning*, **35**(4), 792–815. https://doi.org/10.1080/09588221.2020.1750429

Hua, C., & **Wang, J.** (2023). Virtual reality-assisted language learning: A follow-up review (2018–2022). *Frontiers in Psychology*, **14**, 1153642. https://doi.org/10.3389/fpsyg.2023.1153642

Huizeling, E., Alday, P. M., Peeters, D., & **Hagoort, P.** (2023). Combining EEG and 3D-eye-tracking to study the prediction of upcoming speech in naturalistic virtual environments: A proof of principle. *Neuropsychologia*, **191**, 108730. https://doi.org/10.1016/j.neuropsychologia.2023.108730

Huizeling, E., Peeters, D., & **Hagoort, P.** (2022). Prediction of upcoming speech under fluent and disfluent conditions: Eye tracking evidence from immersive virtual reality. *Language, Cognition and Neuroscience*, **37**(4), 481–508. https://doi.org/10.1080/23273798.2021.1994621

Jiao, L., Lin, Y., Schwieter, J. W., & **Liu, C.** (2026). Learning novel words in an immersive virtual-reality context: Tracking lexicalization through behavioral and event-related-potential measures. *Language Learning*, **76**(1), 5–27. https://doi.org/10.1111/lang.12707

Jiao, L., Zhu, M., Xu, Z., Zhou, G., Schwieter, J. W., & **Liu, C.** (2024). An ERP study on novel word learning in an immersive virtual reality context. *Bilingualism: Language and Cognition*, **27**(1), 128–136. https://doi.org/10.1017/S136672892300038X

Jones, Ō. P., Green, D. W., Grogan, A., Pliatsikas, C., Filippopolitis, K., Ali, N., Lee, H. L., Ramsden, S., Gazarian, K., Prejawa, S., Price, C. J., & **Seghier, M. L.** (2012). Where, when and why brain activation differs for bilinguals and monolinguals during picture naming and reading aloud. *Cerebral Cortex*, **22**(4), 892–902. https://doi.org/10.1093/cercor/bhr161

Kepinska, O., Pereda, E., Caspers, J., & **Schiller, N. O.** (2017). Neural oscillatory mechanisms during novel grammar learning underlying language analytical abilities. *Brain and Language*, **175**, 99–110. https://doi.org/10.1016/j.bandl.2017.10.003

Krejtz, K., Szmidt, T., Duchowski, A. T., & **Krejtz, I.** (2014). Entropy-based statistical analysis of eye movement transitions. In *Proceedings of the symposium on eye tracking research and applications (ETRA '14)* (pp. 159–166). Association for Computing Machinery. https://doi.org/10.1145/2578153.2578176

Lan, J., Fang, Y., Legault, J., & **Li, P.** (2015). Second language acquisition of mandarin Chinese vocabulary: Context of learning effects. *Educational Technology Research and Development*, **63**(5), 671–690. https://doi.org/10.1007/s11423-015-9380-y

Lantz, E. (1996). The future of virtual reality: Head mounted displays versus spatially immersive displays (panel). In *Proceedings of the 23rd annual conference on computer graphics and interactive techniques (SIGGRAPH '96)* (pp. 485–486). Association for Computing Machinery. https://doi.org/10.1145/237170.237289

Lawrence, R. J., Wiggins, I. M., Hodgson, J. C., & **Hartley, D. E. H.** (2021). Evaluating cortical responses to speech in children: A functional near-infrared spectroscopy (fNIRS) study. *Hearing Research*, **401**, 108155. https://doi.org/10.1016/j.heares.2020.108155

Legault, J., Fang, S.-Y., Lan, Y.-J., & **Li, P.** (2019b). 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.** (2019a). Immersive virtual reality as an effective tool for second language vocabulary learning. *Language*, **4**(1), 13. https://doi.org/10.3390/languages4010013

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

Li, P., & **Lan, Y.-J.** (2022). Digital language learning (DLL): Insights from behavior, cognition, and the brain. *Bilingualism: Language and Cognition*, **25**(3), 1–18. https://doi.org/10.1017/S1366728921000353

Li, P., Legault, J., Klippel, A., & **Zhao, J.** (2020). Virtual reality for student learning: Understanding individual differences. *Human Behaviour and Brain*, 28–36. https://doi.org/10.37716/HBAB.2020010105

Lin, Y., Schwieter, J. W., Jiao, L., & **Liu, C.** (2026). Neural tracking of newly learned word lexicalisation in virtual reality: The effects of emotional prosody. *Language, Cognition and Neuroscience*, **41**(5), 579–591. https://doi.org/10.1080/23273798.2026.2650100

Ma, X., Liu, Y., Clariana, R., Gu, C., & **Li, P.** (2023). From eye movements to scanpath networks: A method for studying individual differences in expository text reading. *Behavior Research Methods*, **55**(2), 730–750. https://doi.org/10.3758/s13428-022-01842-3

Macaluso, E., Driver, J., & **Frith, C. D.** (2003). Multimodal spatial representations engaged in human parietal cortex during both saccadic and manual spatial orienting. *Current Biology*, **13**(12), 990–999. https://doi.org/10.1016/S0960-9822(03)00377-4

Macedonia, M., Mathias, B., Lehner, A. E., Reiterer, S. M., & **Repetto, C.** (2023). Grasping virtual objects benefits lower aptitude learners' acquisition of foreign language vocabulary. *Educational Psychology Review*, **35**(4), 115. https://doi.org/10.1007/s10648-023-09835-0

Madsen, J., & **Parra, L. C.** (2022). Cognitive processing of a common stimulus synchronizes brains, hearts, and eyes. *PNAS Nexus*, **1**(1), pgac020. https://doi.org/10.1093/pnasnexus/pgac020

Mahon, B., & **Caramazza, A.** (2008). A critical look at the embodied cognition hypothesis and a new proposal for grounding conceptual content. *Journal of Physiology - Paris*, **102**(1-3), 59–70. https://doi.org/10.1016/j.jphysparis.2008.03.004

Mathias, B., & **von Kriegstein, K.** (2023). Enriched learning: Behavior, brain, and computation. *Trends in Cognitive Sciences*, **27**(1), 81–97. https://doi.org/10.1016/j.tics.2022.10.007

Mayer, K. M., Yildiz, I. B., Macedonia, M., & **von Kriegstein, K.** (2015). Visual and motor cortices differentially support the translation of foreign language words. *Current Biology*, **25**(4), 530–535. https://doi.org/10.1016/j.cub.2014.11.068

Meuter, R. F. I., & **Allport, A.** (1999). Bilingual language switching in naming: Asymmetrical costs of language selection. *Journal of Memory and Language*, **40**(1), 25–40. https://doi.org/10.1006/jmla.1998.2602

Misersky, J., Peeters, D., & **Flecken, M.** (2022). The potential of immersive virtual reality for the study of event perception. *Frontiers in Virtual Reality*, **3**, 697934. https://doi.org/10.3389/frvir.2022.697934

Morgan-Short, K., Faretta-Stutenberg, M., Brill-Schuetz, K. A., Carpenter, H., & **Wong, P. C. M.** (2014). Declarative and procedural memory as individual differences in second language acquisition. *Bilingualism: Language and Cognition*, **17**(1), 56–72. https://doi.org/10.1017/S1366728912000715

Paige, C., Haddad, D. D., Piercy, T., Todd, J., Ward, F., Ekblaw, A., & **Newman, D.** (2024). Development and user study of the operational geology in a virtual environment (OGIVE) platform. *Acta Astronautica*, **224**, 17–36. https://doi.org/10.1016/j.actaastro.2024.07.043

Pausch, R., Proffitt, D., & **Williams, G.** (1997). Quantifying immersion in virtual reality. In *Proceedings of the 24th annual conference on computer graphics and interactive techniques* (pp. 13–18). Addison-Wesley Publishing Co.

Peeters, D. (2019). Virtual reality: A game-changing method for the language sciences. *Psychonomic Bulletin & Review*, **26**(3), 894–900. https://doi.org/10.3758/s13423-019-01571-3

Peng, K., Moussavi, Z., Karunakaran, K. D., Borsook, D., Lesage, F., & **Nguyen, D. K.** (2024). iVR-fNIRS: Studying brain functions in a fully

https://doi.org/10.1017/S1366728926101710 Published online by Cambridge University Press

immersive virtual environment. *Neurophotonics*, **11**(2), 020601. https://doi.org/10.1117/1.NPh.11.2.020601

Rogers, V., **Meara, P.**, **Barnett-Legh, T.**, **Curry, C.**, & **Davie, E.** (2017). Examining the LLAMA aptitude tests. *Journal of the European Second Language Association*, **1**(1), 49–60. https://doi.org/10.22599/jesla.24

Rudling, M., **Rumetshofer, T.**, **Nirme, J.**, **Hedenström, A.**, **Karlsson, T.**, & **Rydberg, A.** (2026). Embodied word learning in schools and sustained attention in virtual reality. *npj Science of Learning*, **11**(1), 9. https://doi.org/10.1038/s41539-025-00395-2

Rumetshofer, T., **Langensee, L.**, **Li, P.**, **Zhao, J.**, **Klippel, A.**, **Wennberg, L.**, **Nilsson, M.**, **Sundgren, P. C.**, **Gullberg, M.**, & **Mårtensson, J.** (2026). Impact of technological immersion and sensorimotor engagement on performance and brain plasticity in short-term second language vocabulary training. *Neurobiology of Language*, **7**. https://doi.org/10.1162/NOL.a.238

Saito, K. (2015). Experience effects on the development of late second language learners' oral proficiency. *Language Learning*, **65**(3), 563–595. https://doi.org/10.1111/lang.12120

Saito, K., **Cui, H.**, **Suzukida, Y.**, **Dardon, D. E.**, **Suzuki, Y.**, **Jeong, H.**, **Révész, A.**, **Sugiura, M.**, & **Tierney, A.** (2022). Does domain-general auditory processing uniquely explain the outcomes of second language speech acquisition, even once cognitive and demographic variables are accounted for? *Bilingualism: Language and Cognition*, **25**(5), 856–868. https://doi.org/10.1017/S1366728922000153

Shadiev, R., & **Li, D.** (2023). A review study on eye-tracking technology usage in immersive virtual reality learning environments. *Computers & Education*, **196**, 104681. https://doi.org/10.1016/j.compedu.2022.104681

Slater, M., & **Wilbur, S.** (1997). A framework for immersive virtual environments (FIVE): Speculations on the role of presence in virtual environments. *Presence: Teleoperators and Virtual Environments*, **6**(6), 603–616. https://doi.org/10.1162/pres.1997.6.6.603

Soltanlou, M., **Sitnikova, M. A.**, **Nuerk, H. C.**, & **Dresler, T.** (2018). Applications of functional near-infrared spectroscopy (fNIRS) in studying cognitive development: The case of mathematics and language. *Frontiers in Psychology*, **9**, 277. https://doi.org/10.3389/fpsyg.2018.00277

Taguchi, N. (2022a). Dispreferred speech acts in virtual reality: Analysis of tone choices and hesitations. *System*, **106**, 102793. https://doi.org/10.1016/j.system.2022.102793

Taguchi, N. (2022b). Immersive virtual reality for pragmatics task development. *TESOL Journal*, **56**(1), 308–335. https://doi.org/10.1002/tesq.3070

Titus, A., & **Peeters, D.** (2024). Multilingualism at the market: A pre-registered immersive virtual reality study of bilingual language switching. *Journal of Cognition*, **7**(1), 1–35. https://doi.org/10.5334/joc.359

Tromp, J., **Peeters, D.**, **Meyer, A. S.**, & **Hagoort, P.** (2018). The combined use of virtual reality and EEG to study language processing in naturalistic environments. *Behavior Research Methods*, **50**(2), 862–869. https://doi.org/10.3758/s13428-017-0911-9

Ullman, M. T., & **Lovelett, J. T.** (2018). Implications of the declarative/procedural model for improving second language learning: The role of memory enhancement techniques. *Second Language Research*, **34**(1), 39–65. https://doi.org/10.1177/0267658316675195

Wan, H., **Zhang, J.**, **Suria, A. A.**, **Yao, B.**, **Wang, D.**, **Coady, Y.**, & **Prpa, M.** (2024). Building LLM-based AI agents in social virtual reality. *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*, 1–7. https://doi.org/10.1145/3613905.3651026

Wang, C.-C., **Hung, J. C.**, & **Chen, H.-C.** (2021). How prior knowledge affects visual attention of Japanese mimicry and onomatopoeia and learning outcomes: Evidence from virtual reality eye tracking. *Sustainability*, **13**(19), 11058. https://doi.org/10.3390/su131911058

Wen, S., **Middleton, M.**, **Ping, S.**, **Chawla, N. N.**, **Wu, G.**, **Feest, B. S.**, **Nadri, C.**, **Liu, Y.**, **Kaber, D.**, **Zahabi, M.**, **McMahan, R. P.**, **Castelo, S.**, **McKendrick, R.**, **Qian, J.**, & **Silva, C. T.** (2025). AdaptiveCoPilot: Design and testing of a NeuroAdaptive LLM cockpit guidance system in both novice and expert pilots. In *2025 IEEE conference virtual reality and 3D user interfaces (VR)* (pp. 656–666). https://doi.org/10.1109/VR59515.2025.00088

Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin & Review*, **9**(4), 625–636. https://doi.org/10.3758/BF03196322

Yang, A. Y., **Qi, R.**, **Liu, G.**, **Zhang, J.**, **Chen, Y.**, & **Hsiao, J. H.** (2023). Humans vs. AI in detecting vehicles and humans in driving scenarios. In *In proceedings of the annual meeting of the cognitive science society* (pp. 45)

Yang, F., **Qian, J.**, **Novotny, J.**, **Badre, D.**, **Jackson, C. D.**, & **Laidlaw, D. H.** (2021). A virtual reality memory palace variant aids knowledge retrieval from scholarly articles. *IEEE Transactions on Visualization and Computer Graphics*, **27**(12), 4359–4373. https://doi.org/10.1109/TVCG.2020.3009003

Yu, S.-Y., **Gao, Z.**, **Zhao, J.**, **Yu, S.-F.**, & **Li, P.** (2026). Immersive virtual reality for second language learning. In P. Li, W. Feng, & Y. Yao (Eds.), *Interdisciplinary research in language sciences: Yuen-ren Chao's legacy* (pp. 169–194). The Hong Kong Polytechnic University Press.

Yu, Y.-T., & **Tsuei, M.** (2023). The effects of digital game-based learning on children's Chinese language learning, attention and self-efficacy. *Interactive Learning Environments*, **31**(10), 6113–6132. https://doi.org/10.1080/10494820.2022.2028855

Yücel, M. A., **Lühmann, A. V.**, **Scholkmann, F.**, **Gervain, J.**, **Dan, I.**, **Ayaz, H.**, **Boas, D.**, **Cooper, R. J.**, **Culver, J.**, **Elwell, C. E.**, **Eggebrecht, A.**, **Franceschini, M. A.**, **Grova, C.**, **Homae, F.**, **Lesage, F.**, **Obrig, H.**, **Tachtsidis, I.**, **Tak, S.**, **Tong, Y.**, … **Wolf, M.** (2021). Best practices for fNIRS publications. *Neurophotonics*, **8**(01). https://doi.org/10.1117/1.NPh.8.1.012101

Zappa, A., **Bolger, D.**, **Pergandi, J.-M.**, **Fargier, R.**, **Mestre, D.**, & **Frenck-Mestre, C.** (2024). The neural correlates of embodied L2 learning: Does embodied L2 verb learning affect representation and retention? *Neurobiology of Language*, **5**(2), 360–384. https://doi.org/10.1162/nol_a_00132

Zappa, A., **Bolger, D.**, **Pergandi, J.-M.**, **Mallet, P.**, **Dubarry, A.-S.**, **Mestre, D.**, & **Frenck-Mestre, C.** (2019). Motor resonance during linguistic processing as shown by EEG in a naturalistic VR environment. *Brain and Cognition*, **134**, 44–57. https://doi.org/10.1016/j.bandc.2019.05.003

Zappa, A., & **Frenck-Mestre, C.** (2023). Embodied second language processing and learning from a neurocognitive perspective. In K. Morgan-Short & J. G. Van Hell (Eds.), *The Routledge handbook of second language acquisition and neurolinguistics* (1st ed., pp. 381–393). New York: Routledge. https://doi.org/10.4324/9781003190912-35

Zappa, A., **Slater, M.**, & **Rodriguez-Fornells, A.** (2025). Social interaction shapes and boosts second language learning: Virtual reality can show us how. *npj Science of Learning*, **10**(1), 90. https://doi.org/10.1038/s41539-025-00381-8

Zhang, Y., **Shi, Z.**, **Liu, Y.**, & **Guo, J.** (2025). The facilitative mechanism of social interaction scenarios on second language vocabulary learning: Evidence from near-infrared spectroscopy hyper-scanning. *Cerebral Cortex*, **35**(4). https://doi.org/10.1093/cercor/bhaf070

Zhao, J., **LaFemina, P.**, **Carr, J.**, **Sajjadi, P.**, **Wallgrün, J. O.**, & **Klippel, A.** (2020). Learning in the field: Comparison of desktop, immersive virtual reality, and actual field trips for place-based STEM education. In *2020 IEEE conference on virtual reality and 3D user interfaces (VR)* (pp. 893–902). IEEE. https://doi.org/10.1109/VR46266.2020.00012

Zhao, J., **Ma, X.**, **Simpson, M.**, **Sajjadi, P.**, **Wallgrün, J. O.**, & **Klippel, A.** (2022). Reference frames and geographic scale: Understanding their relationship in environmental learning. *Cartography and Geographic Information Science*, **49**(1), 70–84. https://doi.org/10.1080/15230406.2021.1942219

Zhao, J., **Simpson, M.**, **Sajjadi, P.**, **Wallgrün, J. O.**, **Li, P.**, **Bagher, M. M.**, **Oprean, D.**, **Padilla, L.**, & **Klippel, A.** (2021). CrowdXR: Pitfalls and potentials of experiments with remote participants. In M. Marchal, J. Ventura, A.-H. Olivier, L. Wang, & R. Radkowski (Eds.), *Proceedings of the 2021 IEEE international symposium on mixed and augmented reality (ISMAR)* (pp. 450–459). IEEE. https://doi.org/10.1109/ISMAR52148.2021.00062

Zhao, J., **Wallgrün, J. O.**, **LaFemina, P. C.**, **Normandeau, J.**, & **Klippel, A.** (2019). Harnessing the power of immersive virtual reality: Visualization and analysis of 3D earth science data sets. *Geo-spatial Information Science*, **22**(4), 237–250. https://doi.org/10.1080/10095020.2019.1621544

Zou, L., **Zhang, Z.**, **Mavilidi, M.**, **Chen, Y.**, **Herold, F.**, **Ouwehand, K.**, & **Paas, F.** (2025). The synergy of embodied cognition and cognitive load theory for optimized learning. *Nature Human Behaviour*, **9**(5), 877–885. https://doi.org/10.1038/s41562-025-02152-2

https://doi.org/10.1017/S1366728926101710 Published online by Cambridge University Press