

Instructional Design for Co-Created and Context-Based VR Chinese Learning Environments

Satoko Sugie 

Hokkai-Gakuen University, Japan

Abstract

The integration of VR and foreign language education offers new opportunities for contextual, interactive, and co-created learning (Durak & Cankaya, 2022). This study investigates how VR enhances co-creative CFL (Chinese as a Foreign Language) at a Japanese university, where students begin with limited language proficiency. The research focuses on the communicative potential of multimodal resources in learner-centered collaboration. Prior research has shown that virtual spaces can increase learner motivation and identity engagement even among struggling learners of English (Chen & Kent, 2019). However, their application in non-English foreign language remains underexplored. We developed VR spaces using “Spatial” through collaboration between students and the teacher. Scenarios were designed to reflect learners’ cultural interests and daily experiences. Classroom sessions incorporated VR-based tasks emphasizing spatial interaction. Data was collected through participant observation and open-ended surveys with 40 students. The results revealed that learners actively co-constructed meaning beyond spoken language, utilizing avatars, shared visual references, and spatial positioning. Motivation increased notably when scenarios aligned with learners’ lived experiences. These findings align with multimodal design approaches in foreign language education, highlighting the importance of integrating linguistic and non-linguistic resources. However, technical barriers such as device access and interface complexity were noted. This study suggests that VR-supported environments enable learners to integrate language with multiple communicative resources and build context-rich, authentic experiences. Such environments shift the focus from traditional four-skill acquisition toward contextually optimized communication.

E-mail: s_sugie@hgu.jp

Keywords: VR, metaverse, CFL (Chinese as a Foreign Language), instructional design, context-based learning, action research

Introduction

Traditional CFL instruction in Japan often emphasizes vocabulary and grammar drills over authentic communication, making it difficult for learners to perceive relevance or purpose in their studies. With the development of immersive technologies such as VR and the metaverse, new opportunities have emerged to reframe CFL learning environments. These tools allow learners to engage in spatial and embodied interaction, creating contexts where language can be experienced and constructed meaningfully (Sugie, 2025). This study explores how VR-supported task-based instruction can address these pedagogical gaps by fostering learner agency, contextual relevance, and multimodal communication.

This action research examines the following research questions:

- RQ1: How do CFL learners perceive the value of VR-based tasks in terms of meaningful learning?
- RQ2: What kinds of learning experiences, needs, and expectations arise through immersive task participation?

Literature Review

This study refers to the ARCS model of motivational design (Keller, 1983) and Self-Determination Theory, which supports autonomous learning (Deci & Ryan, 2009), as its analytical framework. The ARCS model is employed to organize the motivational elements of the VR tasks (Attention, Relevance, Confidence, Satisfaction). Self-Determination Theory is used to confirm how the fulfillment of learners' psychological needs for "Autonomy, Competence, and Relatedness" enhances their intrinsic motivation. From the perspectives of two theories, we explore the learning effects from multiple angles. The analysis of the course evaluation responses was structured around the core constructs of these two theories, using their key elements as analytical dimensions.

Methodology

The study involved 40 first-year students enrolled in a beginner-level CFL course in Fall 2024 at a private Japanese university. Most participants were

humanities majors with limited prior exposure to Chinese. Classes were held weekly for 90 minutes, over 15 sessions.

The instructional design was based on a blended learning course design, integrating both traditional and technology-enhanced components:

1. Textbook-based instruction for vocabulary and grammar, following a standard syllabus;
2. Pronunciation practice via LMS tools using automatic speech recognition to promote self-monitoring;
3. Worksheet-based activities focused on output, such as sentence construction and Q&A practice; and
4. Immersive VR tasks aligned with grammatical objectives, allowing students to apply their learning in simulated social scenarios.

Three original VR tasks were developed collaboratively by the teacher and students using platforms such as Spatial (for VR spaces), Tinkercad and Blender (for 3D modeling), and Monster Mash (for simple interactive objects). These tasks were as follows:

1. Party Preparation: Planning and shopping in Chinese.
2. World Travel, Photo Shooting, and Memory Presentation: Forming travel teams, choosing destinations, taking photos of memorable moments, and reporting travel experiences in Chinese.
3. Room Tidying: Giving and following spatial directions using directional complements.

Each task was aligned with specific grammar points such as “了” (completed action), “要” and “想” (desire and intention), and 趋向补语 (directional complements). Pre-task worksheets and post-task reflections were used to reinforce understanding.

The core of this instructional design is its process of co-creation by the teacher and students. Students are not mere consumers of resources provided by the teacher. Instead, the VR space creation and class activities are built upon student contributions, including their ideas, research, data collection, and added materials. Before the implementation of the classes, some students participated in this project as an extracurricular activity. Students were involved in the following phases:

1. Proposing the theme of “World Travel”: This included devising example sentences and dialogues for countries and regions they wished to visit, using previously learned vocabulary.

2. Conceptualizing and sourcing 3D assets: They searched for typical symbols or iconic landmarks for their travel destinations and collected the corresponding 3D assets.
3. Photographing the travel experience: They took photos to record their travel experiences during the VR trip.
4. Completing the gallery: The final step was to complete a gallery space by displaying their travel photos.

Findings

Learner responses collected through post-task and end-of-semester surveys indicated strong interest and positive emotional engagement with VR-based learning. Over 90% of respondents reported higher motivation and enjoyment. Most learners also agreed that VR tasks deepened their understanding of textbook content.

Qualitative responses emphasized that VR offered “realistic” and “fun” contexts for applying Chinese, lowering anxiety and encouraging active participation. Learners also valued being able to “move and speak naturally” with peers in a virtual space. However, several noted technical issues—including lag, device limitations, and operational difficulties—as barriers to smooth participation.

In terms of instructional needs, many students still desired face-to-face support for pronunciation correction and grammatical guidance. This points to the complementary role of human instruction even in tech-enhanced settings.

Discussion

VR facilitated active, student-centered learning by providing immersive and playful environments where learners could practice Chinese meaningfully. The findings support the relevance of ARCS model (Keller, 1983) and self-determination theory (Deci & Ryan, 2009) in explaining increased motivation.



Figure 1. VR spaces of world travel, photo shooting, and memory presentation.

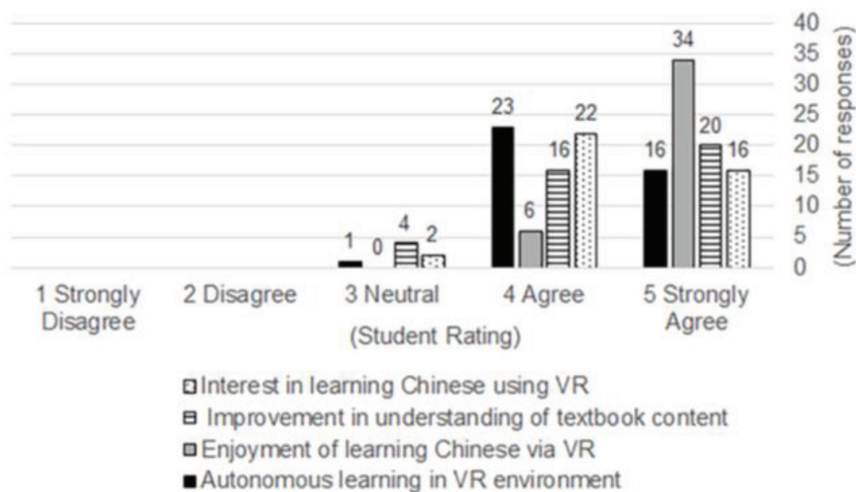


Figure 2. Learners' feedback regarding VR tasks.

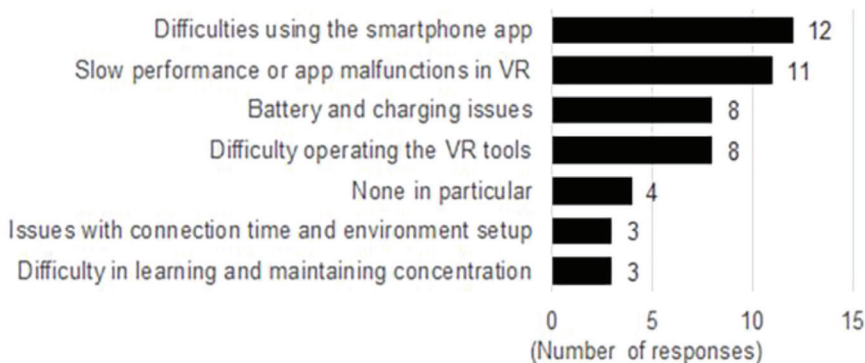


Figure 3. Technical issues and barriers related to VR-based learning.

Three key implications emerged:

1. Task Authenticity: Linking VR tasks to textbook grammar enabled students to apply language meaningfully, turning abstract rules into real-world use.
2. Hybrid Design: A blend of VR and face-to-face interaction proved essential. While VR stimulated engagement, learners still required in-person correction and support.
3. Technical Readiness: Infrastructure and digital literacy remain challenges. Instructors must scaffold VR participation and offer alternatives for learners struggling with devices.

Conclusion

This study demonstrates that VR-enhanced instruction can transform CFL learning into an immersive, co-creative process that fosters learner autonomy and engagement. However, effective implementation depends on balanced instructional design, technological support, and the strategic integration of face-to-face feedback. Future research should explore AI-based feedback systems, multimodal assessment methods, and the scalability of VR-enhanced instruction across diverse educational settings.

References

- Chen, J., & Kent, S. (2019). Task engagement, learner motivation, and avatar identities of struggling English language learners in the 3D virtual world. *System*, 87, 102168. <https://doi.org/10.1016/j.system.2019.102168>
- Deci, E. L., & Ryan, R. M. (2009). Self-determination theory: A macrotheory of human motivation, development, and health. *Canadian Psychology/Psychologie Canadienne*, 49(3), 182–185.
- Durak, G., & Cankaya, S. (Eds.). (2022). *Shaping the future of online learning: Education in the metaverse*. IGI Global. <https://doi.org/10.4018/978-1-6684-6513-4>
- Keller, J. M. (1983). Motivational design of instruction. In C. M. Reigeluth (Eds.), *Instructional design theories and models: An overview of their current status* (pp. 383–433). Lawrence Erlbaum Associates.
- Sugie, S. (2025). Design of multimodal foreign language education. In Proceedings of Language Education Expo 2025. [『マルチモーダルな外国語教育のデザイン』『言語教育エキスポ 2025 予稿集+』]