International Journal of Latest Research in Humanities and Social Science (IJLRHSS)
Volume 07 - Issue 06, 2024
www.ijlrhss.com || PP. 70-88

A Study on Integrating Virtual Reality into Japanese Vocabulary and Kana Syllabary Learning

Ching-Jung Liao*, Zhi-Jie Yang
Department of Information Management, Chung Yuan Christian University,
No. 200, Zhongbei Rd., Zhongli Dist., Taoyuan City 320314, Taiwan
**Corresponding Author*

Abstract: Japan is the most popular travel destination for Taiwanese people, which has led to an increasing prevalence of Japanese language learning. This study aims to provide learners with an immersive learning experience through VR, utilizing speech and object interaction to enhance the understanding and memory of Japanese vocabulary and kana syllabary. This method aims to help learners better grasp foreign vocabulary and improve learning outcomes. Initially, a VR system called JLMVR, set in a museum scenario, was constructed for learning Japanese vocabulary and kana syllabary. The experiment involved five Taiwanese university students with no prior Japanese learning experience using JLMVR, with their learning process recorded. The research method employed lag sequential analysis to encode and analyze the video footage of the learning process, identifying each learner's learning patterns. Japanese proficiency tests were also conducted to examine the impact of JLMVR on beginners' learning of Japanese vocabulary and kana syllabary. The research findings are as follows:

1. The scores of three participants significantly improved after learning through VR, while the other two participants showed no obvious improvement, indicating the influence of individual differences on learning outcomes.
2. VR provides a highly immersive learning experience and has a significant effect on enhancing learners' motivation and memory depth.
3. The study found that an appropriate level of challenge is key to maintaining a state of flow, and the difficulty of the materials should be adjusted according to individual abilities.

In conclusion, the study's results demonstrate that VR technology exhibits unique advantages in language learning and has the potential to become an important tool in this field. Future research could further explore the adaptability of learners from different backgrounds and optimize interaction design in VR to enhance its effectiveness in language learning.

Keywords: Virtual Reality, Japanese vocabulary, Japanese kana syllabary, Hiragana, Lag sequential analysis.

1. Introduction

COVID-19 has significantly impacted the global tourism industry. However, with the development of virtual reality (VR) technology, individuals can still experience the scenic beauty of various global destinations from home during the pandemic, providing an alternative means of travel. Sarkady et al. (2021) found that when physical travel is restricted due to external factors such as the COVID-19 pandemic, the use of VR for travel experiences increases people's desire to travel. Japan is the most popular travel destination for Taiwanese people, prompting many to learn Japanese to facilitate communication with locals. Traditionally, Japanese learning has relied on textbooks or printed materials, and with technological advancements, many online language courses have emerged. Despite this, many foreign language learners experience anxiety during speaking exercises or when using the language, and traditional classroom instruction often fails to provide sufficient resources to overcome these issues (Harris & Reid, 2005; Horwitz et al., 1986; Liu & Jackson, 2008).

Advances in VR have introduced more interactive and immersive learning experiences. Compared to other learning methods, VR offers a more engaging learning environment where learners can interact directly with elements within the virtual setting. VR also provides the capability for error correction without the embarrassment that may arise in real-life interactions, thereby reducing learners' anxiety. VR’s ability to create an immersive, personalized environment allows learners to make mistakes without fear of embarrassment or anxiety, encouraging more active participation. Additionally, learning through practical engagement in VR can enhance memory retention, with recall rates reaching up to 90%, highlighting the importance of VR applications in education (Dale, 1969; Harris & Reid, 2005).

Nevertheless, Villena-Taranilla et al. (2022) noted that research on VR in language learning remains relatively limited. Parmaxi (2023) pointed out that most VR language learning studies focus on English, with relatively few studies on Japanese. The foundation of Japanese language learning is kana syllabary, which is crucial for pronunciation. Mastery of kana syllabary enables learners to pronounce words upon seeing them, forming a vital foundation in Japanese language education. Learners who are familiar with kana syllabary can confidently read aloud and improve their overall communication and learning outcomes. However, most learners rely on rote memorization to learn kana syllabary. This study employs VR as a learning tool, focusing on Japanese vocabulary and kana syllabary, to provide learners with an immersive environment that offers an alternative to traditional learning methods and examines the effectiveness of this emerging technology in language acquisition.

2. Research Purpose and Questions

Based on the research background and motivation, this study constructs a VR system for learning Japanese vocabulary and syllabary. This system combines kana syllabary and daily objects in an immersive virtual environment to explore the learning patterns of beginners and the effectiveness of VR in Japanese language acquisition. The research objectives are as follows:

1. To identify the learning patterns of beginners using the JLMVR for Japanese vocabulary and syllabary learning.
2. To understand the effectiveness of the JLMVR on the learning outcomes of Japanese beginners.

Accordingly, the research questions are:

1. What are the learning patterns of Japanese beginners using the JLMVR for vocabulary and syllabary learning?
2. What is the impact of JLMVR on the learning outcomes of Japanese beginners?

3. Literature Review

This study examines the application of VR in language learning, e-learning in Japanese education, and related research on lag sequential analysis. The review is divided into sections covering virtual reality, e-learning applications in Japanese, VR applications in language learning, VR learning theories, and lag sequential analysis, each explained in detail.

3.1 Virtual Reality

VRis a simulated environment created through computer technology, immersing users in a three-dimensional virtual world via specialized head-mounted displays or multi-projection systems. As early as 1965, Sutherland introduced the concept of "The Ultimate Display," describing interactive graphics and feedback devices, pioneering the field of computer graphics. In 1968, Sutherland further detailed a head-mounted display, elucidating the basic principles and ideas behind three-dimensional display technology, marking an early form of VR (Sutherland, 1968). The term "virtual reality" was first coined in 1986 by Jaron Lanier, founder of the Visual Programming Languages (VPL) research institute, which was also the first company to sell VR glasses. Since the emergence of the term, the concept of VR has continually evolved. Heim (2000) defined VR as a method of replacing primary sensory input with computer-generated data, creating an illusion of being in another place. Yoh (2001) expanded this definition, stating that VR is not merely limited to material or hardware, nor just a mental state or imagination, but is also a tool for conveying content, a constructive and abstract medium aligning with Sutherland's idea of the "ultimate display."

VR is considered a hybrid term that combines technology, individual experience, and social relationships. Key concepts in VR include immersion, interactivity, and presence (Hillis, 1999). Sherman and Craig (2003) argued that VR can perceive user states and operations, substituting or enhancing sensory feedback in ways that make users feel immersed in the simulation. They identified four elements of VR: virtual environment, virtual presence, sensory feedback, and interactivity, consistent with Hillis (1999). Hence, VR is not just a technology but an essential tool for altering perception and interaction, enabling experiences akin to the real world.

3.2 e-Learning Applications in Japanese

Language learning is crucial in cross-cultural communication, with different learning environments and media significantly affecting learning outcomes. Researchers strive to understand how various learning methods impact students to enhance language learning efficacy. Oxford et al. (1993) identified motivation as the best predictor of success in language learning. To boost motivation for learning Japanese, Udjaja et al. (2019) developed a role-playing game on the Nintendo DS lite (NDSL), utilizing the touch panel for writing Japanese characters, allowing players to learn Japanese through gameplay. With the proliferation of smartphones, researchers shifted focus to mobile learning. Kurniawan & Novita (2019) developed "AKEBI," an Android-based app for learning Japanese characters systematically, emphasizing character levels, stroke order, and pronunciation, making it easier for learners to understand and remember.

Augmented Reality (AR) and VR have emerged as technologies that can bring learning closer to real-life experiences. Early scholars defined AR by its ability to combine the real and virtual, provide real-time interaction, and use 3D positioning, which can immerse learners in a new learning space and enhance motivation (Azuma, 1997; Georgiou & Kyza, 2018). In Japanese learning, Geng & Yamada (2020) developed an AR system for learning compound verbs, allowing interaction with virtual characters to enhance motivation. The distinction between VR and AR lies in VR creating an immersive environment for interaction with various components (Farshid et al., 2018; Ye, 2022).

Different e-learning methods in Japanese education yield various results, with motivation being crucial in language learning. AR and VR have the potential to enhance learning motivation, especially in providing immersive learning experiences. These technologies allow learners to study languages in more interactive and realistic environments, thereby improving learning efficiency and interest.

3.3 Applications of VR in Language Learning

The rapid development of technology has revolutionized education, with VR providing an interactive and engaging learning environment beyond traditional methods. VR has been shown to increase learner motivation (R. Liu et al., 2022; Su et al., 2022). In virtual spaces, learners can experience interactive activities unavailable in traditional education and explore freely without real-world risks, aiding in the comprehension and memory of complex or abstract concepts. However, successful technological applications require advanced technical support and well-designed content that neither overwhelms learners with difficulty nor bores them with simplicity, facilitating flow states (Csikszentmihalyi, 1991; Suh & Prophet, 2018).

Abich IV et al. (2021) reviewed VR applications across various fields until 2019, finding significant use in medicine for high-risk surgeries, training healthcare personnel, and treating conditions like arachnophobia (Garcia-Palacios et al., 2002). VR is also valuable in disaster training, enabling users to learn safety measures in simulated scenarios (Li et al., 2017). Villena-Taranilla et al. (2022) noted that while VR research in language learning is limited, immersive VR positively impacts declarative knowledge learning. Language learning faces obstacles such as the quality of the learning environment and learner motivation. VR’s interactive, realistic environments can alleviate language learning anxiety and improve engagement (Zheng, 2008; Al Farsi et al., 2021).

Researchers suggest using VR to create immersive learning environments to reduce language learning anxiety. Combining AI and multi-agent technology in immersive VR learning systems, such as AI-based chatbots in the metaverse, can enhance interest, practical engagement, and social interaction (Yu, 2023; Rajesh et al., 2023). VR can also overcome communication and cultural barriers, benefiting cross-cultural and multilingual communication skills (Ciupe et al., 2023; Ader et al., 2023). Integrating VR and AI in online English-speaking courses has improved language skills (Richard & Rajakumari, 2023; Li & Qin, 2023).

VR’s advantages include providing highly interactive and realistic learning environments, enhancing language learners’ memory and reducing anxiety. These features make VR an effective tool for increasing motivation and participation in language learning. Thus, using VR can create a more engaging and effective Japanese learning environment, improving learning outcomes and overall learner experience.

Integrating VR into language learning provides a novel interactive experience, breaking the limitations of traditional teaching by creating realistic and vivid learning environments. VR allows learners to immerse in various simulated contexts, enhancing language learning. Researchers have developed VR environments and applied different learning methods, such as "searching for items" to learn new language knowledge, enhancing listening comprehension and vocabulary skills (Garcia et al., 2019; Jia & Liu, 2019). VR has been used to teach contextually, such as learning Japanese communication through simulated convenience store or izakaya conversations, improving speaking skills (ITO, 2022; Liu & Zou, 2022).

Although VR learning systems have shown significant effectiveness, most are limited to English or use Romanized versions for pronunciation. For languages like Japanese, incorporating syllabary systems is necessary.

3.4 VR Learning Theories

Integrating technology into learning is a significant trend, combining traditional learning theories with advanced VR technology to create better learning experiences. Constructivist learning theory, as proposed by Dewey (1916), emphasizes active participation and sensory experience in constructing knowledge. Constructivism in language teaching focuses on action-oriented, collaborative learning, emphasizing learner autonomy and cultural awareness (Suhendi, 2018).

VR offers an interactive, contextual learning environment, promoting active participation and knowledge construction. Constructivism principles are foundational for understanding VR’s role in learning (Cheng & Wang, 2011; Huang & Liaw, 2018). Mayer's multimedia learning theory highlights balancing visual and verbal information to maximize processing capacity (Mayer, 2002). Combining constructivist principles and multimedia learning theory, VR can enhance learner engagement and cognitive depth (Gilakjani, 2012; Teng, 2023).

VR’s immersive experience provides additional advantages, making Daghestani et al.'s (2012) VR learning cognitive theory more applicable than Mayer’s multimedia learning theory due to VR’s interactivity. Csikszentmihalyi's (1991) "flow" state, characterized by deep focus and engagement, can enhance learning experiences and motivation (Erhel & Jamet, 2019). Achieving flow in learning activities positively impacts outcomes, with factors such as clear goals, immediate feedback, and matching challenges with skills being essential (Marty-Dugas et al., 2021; Liu et al., 2022).

Based on these theories, creating immersive VR learning environments is crucial for improving language learning outcomes. Combining constructivism and multimedia learning theory, VR can provide rich, contextual experiences, enhancing learner engagement and cognitive depth. Utilizing VR technology in language learning can effectively boost learning outcomes, especially when course design maximizes presence and facilitates flow states.

3.5 Lag Sequential Analysis

Lag sequential analysis (LSA) referenced in this study was first introduced by Sackett (1974). Bakeman and Gottman (1997), in their book "Observing Interaction", describe how behaviors of research subjects are encoded into events and arranged sequentially to observe a series of event sequences. Using a state transition table, behavior sequence transitions can be compared and analyzed. Researchers can use the analysis results to infer and determine whether the sequential relationships among learning behaviors reach statistical significance. In this study, the term "sequence analysis" refers to lag sequential analysis, a method of encoding that helps researchers understand the sequential patterns of learners' behaviors (Bakeman & Gottman, 1997).

This method determines whether the order of specific learning behavior patterns achieves statistical significance. The implementation of sequence analysis involves encoding each learner's behavior according to specific categories. To ensure rigorous analysis and avoid subjectivity from a single coder, two coders with relevant academic backgrounds must complete the encoding and validate the consistency between them. If discrepancies arise, they must discuss to reach a consensus and revise the encoding. Based on the encoded results, researchers calculate the frequency and proportion distribution of each code and perform sequence analysis. They then create a state transition table, where the vertical columns represent the initial behaviors, and the horizontal rows represent subsequent behaviors. A z-value greater than 1.96 indicates that the sequence is statistically significant.

LSA is commonly used to study or understand the ways and patterns of human-machine interactions. For example, Zarzour et al. (2020) explored the impact of integrating social network functions into an e-book system on student learning processes. Through experiments and LSA, the study analyzed student behavior patterns, revealing significant differences in social interaction behaviors between high and low engagement students. This has important implications for enhancing student engagement and designing learning systems based on social networks. In addition to analyzing human-machine interactions, LSA can also analyze learning behavior patterns among different learners. Tlili et al. (2023) used LSA to investigate how cultural backgrounds influence student learning behavior patterns. LSA also provides a good explanation for knowledge construction behaviors, helping to track learner behavior patterns and identify differences between high-achieving and low-achieving groups. By analyzing the sequence of behaviors during the learning process, researchers can better understand how knowledge is constructed and how students achieve different levels of understanding in their learning (Sun et al., 2021).

In summary, LSA can be used to reveal and understand behavior patterns and their cross-dependencies. This analysis method significantly impacts understanding interactive behaviors and learner behavior patterns in the learning process. In this study, LSA is used to understand learners' behavior patterns when interacting with virtual environments and to identify key factors affecting learning outcomes.

4. System Framework and Research Methodology

4.1 System Framework of JLMVR

This study constructs a Japanese vocabulary and kana syllabary learning system named JLMVR, which integrates virtual reality technology to explore the learning behavior patterns and outcomes of beginners using the system. The system framework of JLMVR is illustrated in Figure 1. The necessary resources and development tools include 3D scenes and objects, MS Azure speech recognition technology, the Unity 3D development engine, Japanese language learning materials, a Chinese translation database, and VR equipment. The system is developed using Unity 3D, with programming in C# to integrate scenes, objects, characters, and user interfaces. An example of a 3D scene is shown in Figure 2, where various learning objects are displayed in a showcase within the VR environment. A phonetic chart of the Japanese syllabary is framed in red on the wall.

When learners take an object from the showcase, they can view its pronunciation and corresponding Hiragana characters. Through an interactive approach that combines physical objects with audio, learners can verbally repeat the pronunciation and immediately verify its accuracy. If encountering unfamiliar vocabulary, learners can move the object to the area where the phonetic chart is located. The chart will then light up according to the Hiragana marking of the object and simultaneously play the pronunciation, enabling learners to clearly identify the Hiragana characters that make up the word. Additionally, each Hiragana character on the phonetic chart can independently produce sound, allowing learners to engage in immersive Japanese language learning within this VR environment.

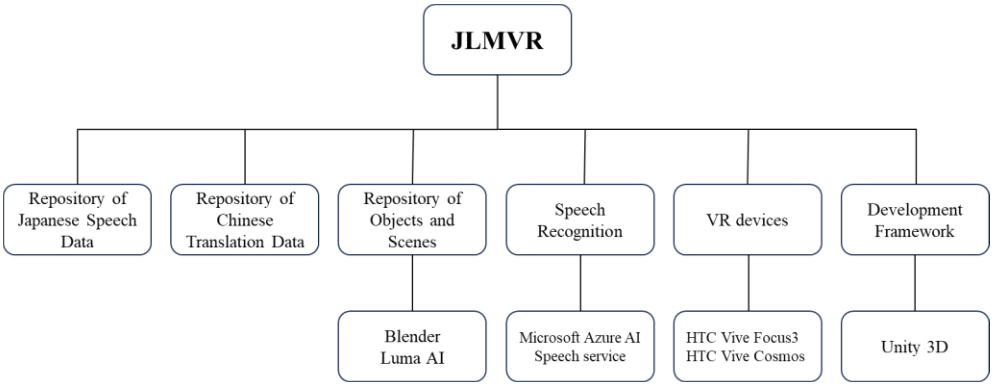

Figure 1: System framework of JLMVR

Figure 2: Examples of JLMVR scenarios

4.2 Research Process

This study recruited five beginners, aged between 20 and 30, who had no prior experience in learning Japanese, as participants for the experiment. During the experimental process, participants were required to engage in learning activities within the JLMVR system three days a week, with each day consisting of two learning sessions. Each session lasted 30 minutes, followed by a 15-minute break, and the experiment spanned two weeks. To meticulously record each participant's behavior, the entire experimental process was videotaped, resulting in a total of 360 minutes of footage per participant. This footage was further segmented at 5-second intervals for detailed analysis, with each segment being coded based on a predefined learning behavior coding scheme.

To conduct sequence analysis, the method and standards for behavior coding in the learning process videos were first established, defining a set of behaviors according to the learning activities. The coding scheme of behaviors in Japanese learning is detailed in Table 1. Coders analyzed the learning process using the predefined behavior content, and upon completing the coding, the results from two coders were integrated. The combined coding results were then converted into an event transition frequency table, which displayed the transition frequencies between different learning behaviors.

The learning process was video-recorded, and two coders performed event coding of the learning behaviors using the Cohen Kappa coefficient to calculate inter-coder reliability (Cohen, 1960). A Kappa coefficient between 0.6 and 0.8 indicated good to excellent consistency and sufficient reliability (Landis & Koch, 1977). If the Kappa coefficient was less than 0.6, the coders adjusted the coding methods.

Table 1: The coding scheme of behaviors in Japanese learning using JLMVR

Code	Phase	Description
C	Controlling VR	The learner performs operations within JLMVR that are not directly related to learning, such as browsing the environment and clicking menus.
Cogn	Cognition	The learner observes objects within JLMVR.
L	Listening	The learner listens to audio within JLMVR.
LH	Learn Hiragana	The learner engages in learning hiragana using a hiragana chart.
LHO	Learn Hiragana with Objects	The learner interacts with objects while learning hiragana in front of a hiragana chart.
R	Repetition	The learner repeats verbal exercises.
S	Speaking	The learner engages in speaking.
T	Thinking	The learner engages in thinking.

5. Experimental Results and Discussions

5.1 Functions of the JLMVR System

This study has developed the JLMVR system, enabling learners to study Japanese vocabulary and the kana syllabary. Some functionalities are illustrated in Figures 3 to 6. When a learner picks up an object, the system displays its name in both Chinese and Hiragana, as shown in Figure 3. By clicking on the "Listen" and "Answer" buttons, learners can access different features: selecting the "Listen" option prompts the system to pronounce the correct reading, while selecting the "Answer" option activates the speech recognition feature, displaying the learner's spoken Japanese word, as shown in Figure 4. For a more detailed understanding of the object's pronunciation, learners can bring the object into contact with the floor in front of the kana syllabary chart, causing the chart to sequentially highlight the corresponding Hiragana characters, as illustrated in Figure 5. If learners need to learn individual Hiragana characters, they can use a teaching stick to touch a character on the wall-mounted kana syllabary chart, causing it to pronounce and flash individually, as depicted in Figure 6.

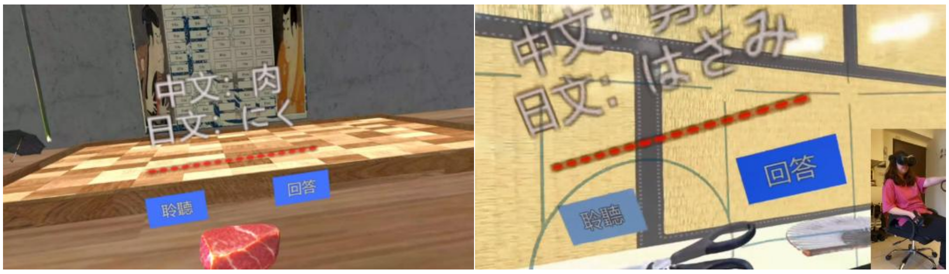

Figure 3: Simultaneous display of Chinese and Japanese for learning objects

Figure 4: Utilizing the "Answer" function to pronounce a learning object

Figure 5: Sequential display of "とうもろこし(corn)" and its pronunciation on the kana syllabary chart

Figure 6: Learners using a teaching stick to touch and illuminate Hiragana characters on the kana syllabary chart

5.2 Analysis of learning behavior patterns using JLMVR for Japanese vocabulary and kana syllabary

The lag sequential analysis used in this study to analyze learning behaviors is based on coding methods comprising device operation, objects within the cognitive system, the learner's acquisition of hiragana, and behaviors such as listening, speaking, repeating, and thinking while learning hiragana through objects. Each participant accumulated a total of 360 minutes of recorded learning videos, which were coded every 5 seconds, resulting in a total of 4,320 behavior codes per participant. With five participants in total, this amounted to 21,600 behavior codes. The coding results from the two coders were integrated and subjected to Kappa consistency comparison, with results of 0.965, 0.948, 0.957, 0.951, and 0.969, respectively.

This study involved five participants, with the results of the sequential analysis described sequentially from Experimenter 1 to Experimenter 5. The frequency and percentage of various codes for Experimenter 1's Japanese learning behaviors using JLMVR are shown in Table 2. The most frequent event was S (Speaking, 1,098 times, 25.4%), indicating that Experimenter 1 is a proactive language learner. Event R (Repeat, 588 times, 13.6%) was also relatively frequent, suggesting that Experimenter 1 prefers to reinforce learning through repetitive practice. Based on the adjusted residuals (Z-value) presented in Table 3, highly significant dual-event transition sequences can be identified. The significant behavior pattern diagram is illustrated in Figure 7, revealing that the patterns LH→T, L→S, T→LH, S→R, and C→L are particularly notable. The significant sequence behavior patterns of Experimenter 1's Japanese learning using JLMVR are shown in Table 5.

Table 2: Frequency and percentage of various codes for Experimenter 1's Japanese learning using JLMVR

Code	Frequency	Percentage %
S	1098	25.4%
L	818	18.9%
R	588	13.6%
LHO	433	10.0%
C	430	10.0%
T	402	9.3%
LH	400	9.3%
Cogn	151	3.5%

Table 3: Adjusted residuals table (Z-value) of Experimenter 1's behaviors during Japanese learningusing JLMVR

Z	C	Cogn	L	LH	LHO	R	S	T
C	-5.964	**2.071**	**7.638**	-1.912	-0.102	-5.825	-0.779	-1.811
Cogn	**4.131**	-1.836	**1.978**	-1.127	-0.839	-2.073	-2.204	0.65
L	-2.582	-1.779	-17.282	-5.433	**2.172**	-3.428	**11.64**	-6.409
LH	1.699	0.144	-3.582	-3.343	-3.402	-3.757	-2.424	**13.843**
LHO	1.545	-1.705	-1.781	1.647	-6.073	0.029	-1.081	**2.076**
R	0.59	-0.405	0.681	-2.175	-1.472	-11.076	**2.08**	-1.056
S	-3.814	-0.594	**2.15**	-4.376	-3.497	**9.894**	-23.151	-3.579
T	-1.9	-1.862	-2.573	**11.474**	1.377	-2.769	-0.288	-7.311

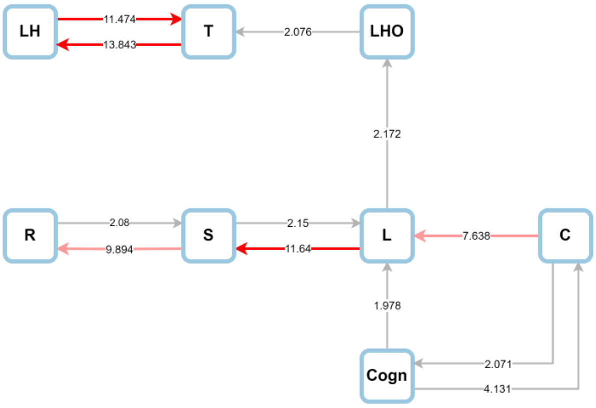

Figure 7: Behavior pattern diagrams of Experimenter 1's Japanese learning using JLMVR

Table 5: Significant sequence behavior patternsof Experimenter 1's Japanese learning using JLMVR

C→ Cogn (Z = 2.071); C→ L (Z = 7.638);
Cogn → C (Z = 4.131); Cogn → L (Z = 1.978)
L→ LHO (Z = 2.172); L→ S (Z = 11.64); LH → T (Z = 13.843)
LHO → T (Z = 2.076); R→ S (Z = 2.08)
S→ L (Z = 2.15); S→ R (Z = 9.894); T→ LH (Z = 11.474)

The frequency and percentage of various codes used by Experimenter 2 in Japanese learning with JLMVR are shown in Table 6. The frequency of event C reached 27% (1167 occurrences), indicating that the experimenter might often engage in activities not directly related to learning, potentially adversely affecting learning outcome. However, the activities for event L (749 occurrences, 17.3%) and event S (705 occurrences, 16.3%) are also frequent, suggesting they might contribute to short-term progress in language learning. The adjusted residuals (Z-value) are shown in Table 7, and the significant behavior pattern diagram is illustrated in Figure 8, revealing highly significant patterns as C→L, S→R, L→S, and LH→C. Significant sequence behavior patterns of Experimenter 2's Japanese learning using JLMVR are shown in Table 8.

Table 6: Frequency and percentage of various codes for Experimenter 2's Japanese learning using JLMVR

Code	Frequency	Percentage %
C	1167	27.0%
L	749	17.3%
S	705	16.3%
Cogn	537	12.4%
LH	519	12.0%
T	223	5.2%
R	216	5.0%
LHO	204	4.7%

Table 7: Adjusted residuals table (Z-value) of Experimenter 2's behaviors during Japanese learning using JLMVR

Z	C	Cogn	L	LH	LHO	R	S	T
C	-11.697	-2.873	**12.177**	-0.719	-0.979	-5.468	-3.055	-3.616
Cogn	**2.306**	-6.314	-2.142	**5.757**	-1.161	-2.333	-2.644	0.439
L	-1.312	-0.923	-15.811	-2.348	**2.866**	-1.979	**6.633**	-4.628
LH	**6.354**	**4.723**	-3.475	-6.66	-2.15	-3.003	-5.509	**3.928**
LHO	0.9	1.833	-2.72	**2.698**	-2.998	-0.896	-1.598	0.866
R	-0.26	-1.711	-0.8	-3.555	-1.606	-4.157	**5.653**	0.791
S	-3.152	-1.847	**2.833**	-4.731	-1.881	**8.934**	-13.958	0.637
T	-1.622	-0.825	-4.421	**3.282**	-0.189	-0.92	**4.861**	-4.005

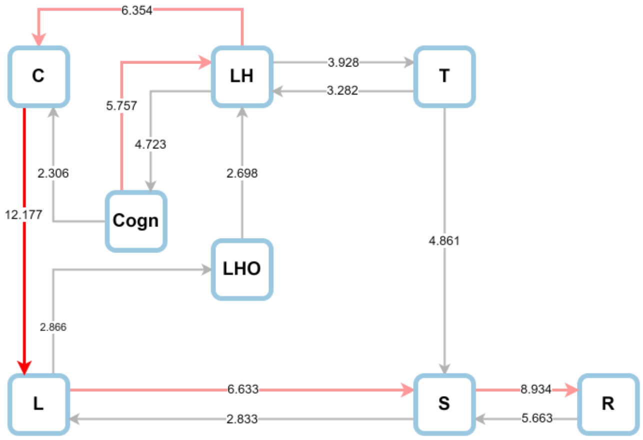

Figure 8: Behavior pattern diagrams of Experimenter 2's Japanese learning using JLMVR

Table 8: Significant sequence behavior patternsof Experimenter 2's Japanese learning using JLMVR

C →L (Z = **12.177**); Cogn →C (Z = 2.306); Cogn →LH (Z = 5.757)
L →LHO (Z = 2.866); L →S (Z = 6.633)
LH→ C(Z = 6.354); LH→ Cogn(Z = 4.723); LH→ T(Z = 3.928)
LHO→ LH(Z = 2.698); R→S(Z = 5.653); S→ L(Z = 2.833); S→ R(Z = 8.934)
T→ LH(Z = 3.282); T→ S(Z = 4.861)

The frequency and percentage of various codes used by Experimenter 3 during Japanese language learning with JLMVR are presented in Table 9. Although the frequency of event L (1706, 39.5%) is the highest among all experimenters and the frequency of event S is also relatively high, the frequency of event R is extremely low. This indicates that while Experimenter 3 tends to pronounce words immediately after hearing them, they lack the action of repetition, which may lead to quicker forgetting of new vocabulary. The adjusted residual values (Z-values) are shown in Table 10, and the significant behavior patterns diagram is illustrated in Figure 9. Notably, the most significant behavior patterns observed are LHO→T, L→S, C→L, and S→L. The significant sequence behavior patterns of Experimenter 3's Japanese learning using JLMVR are detailed in Table 11.

Table 9: Frequency and percentage of various codes for Experimenter 3's Japanese learning using JLMVR

Code	Frequency	Percentage %
L	1706	39.5%
C	941	21.8%
S	631	14.6%
Cogn	621	14.4%
LH	172	4.0%
LHO	170	3.9%
R	45	1.0%
T	34	0.8%

Table 10: Adjusted residuals table (Z-value) of Experimenter 3's behaviors during Japanese learning using JLMVR

Z	C	Cogn	L	LH	LHO	R	S	T
C	-11.123	0.623	**13.443**	**3.298**	-3.779	-3.242	-13.381	-2.454
Cogn	**5.937**	-7.386	**6.778**	1.628	-2.997	-2.177	-10.951	-1.982
L	-7.552	-5.795	-61.602	-5.751	0.955	0.395	**14.655**	-5.15
LH	**4.253**	**2.481**	-1.643	-1.283	-1.163	-1.063	-2.935	0.479
LHO	1.901	-1.922	0.377	-1.207	-2.487	**2.07**	-5.377	**19.424**
R	-0.551	-0.156	**3.232**	-1.068	-0.026	-0.885	-2.974	0.862
S	-6.134	-2.858	**10.681**	-1.836	-3.988	-1.935	-18.026	-2.73
T	-1.186	-1.871	**2.98**	0.478	1.743	-0.668	-1.842	-0.506

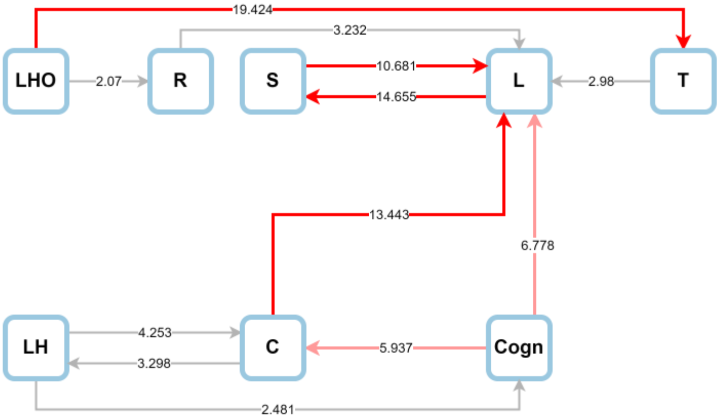

Figure 9: Behavior pattern diagrams of Experimenter 3's Japanese learning using JLMVR

Table 11: Significant sequence behavior patternsof Experimenter 3's Japanese learning using JLMVR

C →L (Z = 13.443); C →LH (Z = 3.298)
Cogn →C (Z = 5.937); Cogn →L (Z = 6.778); L →S (Z = 14.655)
LH →C (Z = 4.253); LH →Cogn (Z = 2.481)
LHO →R (Z = 2.07); LHO →T (Z = **19.424**)
R →L (Z = 3.232); S →L (Z = 10.681); T →L (Z = 2.98)

The frequency and percentage of various codes used by Experimenter 4 during Japanese language learning with JLMVR are presented in Table 12. Events S (1286, 29.8%) and L (1087, 25.2%) together account for more than half of the total, indicating that the learner's language learning style tends towards an outgoing personality, unafraid of the opinions of others. Additionally, the frequency of event LH (743, 17.2%) is the highest among all experimenters, suggesting that Experimenter 4 values fundamental language knowledge and demonstrates a comprehensive learning approach. The adjusted residual values (Z-values) are shown in Table 13, and the significant behavior patterns diagram is illustrated in Figure 10. The most significant behavior patterns observed are LH→T, L→S, T→LH, S→L, C→L, Cogn→C, and R→L. The significant sequence behavior patterns of Experimenter 4's Japanese learning using JLMVR are detailed in Table 14.

Table 12: Frequency and percentage of various codes for Experimenter 4's Japanese learning using JLMVR

Code	Frequency	Percentage %
S	1286	29.8%
L	1087	25.2%
LH	743	17.2%
T	371	8.6%
C	368	8.5%
R	221	5.1%
Cogn	209	4.8%
LHO	35	0.8%

Table 13: Adjusted residuals table (Z-value) of Experimenter 4's behaviors during Japanese learning using JLMVR

Z	C	Cogn	L	LH	LHO	R	S	T
C	-4.232	**5.685**	**8.526**	-1.301	-1.339	-3.621	-5.849	-2.752
Cogn	**7.46**	-2.044	**4.439**	-1.697	-0.922	-2.871	-3.618	-3.111
L	-3.47	-2.016	-31.831	-10.627	-1.809	-1.864	**18.62**	-9.423
LH	-1.862	-1.194	-8.108	-5.8	0.544	-2.73	-1.785	**18.641**
LHO	-1.293	0.242	1.037	0.709	-0.41	-0.428	-1.81	1.754
R	-2.21	-1.393	**7.261**	-1.163	**1.97**	-4.106	-4.976	-0.119
S	-1.085	-3.003	**8.729**	-4.232	-0.769	0.538	-33.433	-5.417
T	-3.555	-2.677	-7.256	**16.249**	-0.453	-1.233	0.74	-7.124

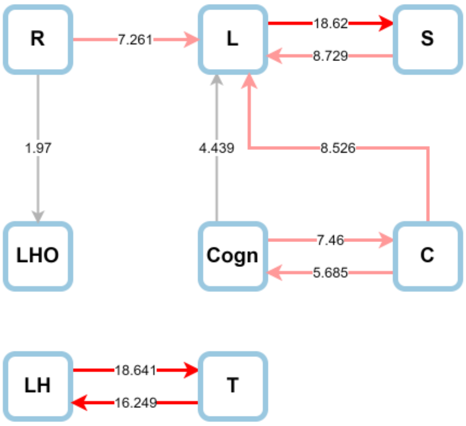

Figure 10: Behavior pattern diagrams of Experimenter 4's Japanese learning using JLMVR

Table 14: Significant sequence behavior patternsof Experimenter 4's Japanese learning using JLMVR

C → Cogn (Z = 5.685); C → L (Z = 8.526); Cogn → C (Z = 7.46); Cogn → L (Z = 4.439);
L→S (Z = 18.62); LH→T (Z = **18.641**); R→L (Z = 7.261); R→LHO (Z = 1.97)
S→L (Z = 8.729); T→LH (Z = 16.249);

The frequency and percentage of various codes used by Experimenter 5 during Japanese language learning with JLMVR are presented in Table 15. Events S (977, 22.6%) and LH (926, 21.4%) occur with higher frequency, while the frequency of event LHO (645, 14.9%) is also relatively high. This indicates that the learner frequently interacts with objects to learn hiragana, suggesting that Experimenter 5 tends to adopt a step-by-step strategy for learning. The adjusted residual values (Z-values) are shown in Table 16, and the significant behavior patterns diagram is illustrated in Figure 11. The most significant behavior patterns observed are L→S, LH→T, T →LH, Cogn→L, and LHO→LH. The significant sequence behavior patterns of Experimenter 5's Japanese learning using JLMVR are detailed in Table 17.

Table 15: Frequency and percentage of various codes for Experimenter 5's Japanese learning using JLMVR

Code	Frequency	Percentage %
S	977	22.6%
LH	926	21.4%
L	721	16.7%
LHO	642	14.9%
T	321	7.4%
Cogn	303	7.0%
R	258	6.0%
C	172	4.0%

Table 16: Adjusted residuals table (Z-value) of Experimenter 5's behaviors during Japanese learning using JLMVR

Z	C	Cogn	L	LH	LHO	R	S	T
C	-1.937	-0.561	**4.013**	1.253	-0.936	-1.904	-0.602	-3.139
Cogn	1.839	-3.669	**6.904**	-3.544	-0.542	-2.251	-0.296	-3.549
L	-1.472	0.627	-17.687	-8.135	-6.849	**3.53**	**15.635**	-7.752
LH	0.886	-3.969	-4.63	-10.55	**5.089**	-4.243	-3.037	**12.096**
LHO	0.027	-3.495	-1.988	**6.682**	-9.417	-4.616	-1.91	**4.767**
R	-2.001	1.262	**4.887**	-0.687	**2.014**	-4.829	-2.485	-4.78
S	-1.8	**3.063**	**2.867**	-3.864	-0.324	**3.783**	-21.655	-3.974
T	-1.467	-4.142	-2.326	**11.322**	0.14	-4.086	-1.794	-5.231

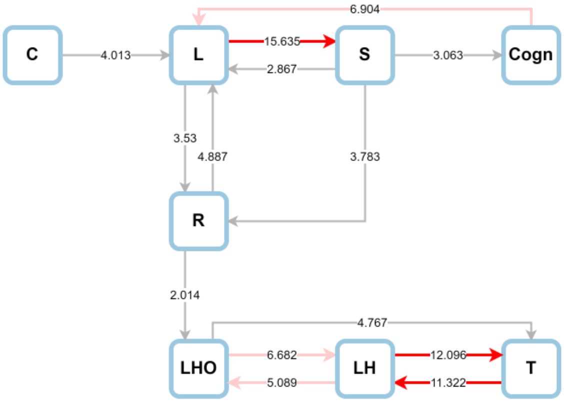

Figure 11: Behavior pattern diagrams of Experimenter 5's Japanese learning using JLMVR

Table 17: Significant sequence behavior patternsof Experimenter 5's Japanese learning using JLMVR

C→L(Z = 4.013); Cogn→L (Z = 6.904); L→R (Z = 3.53); L→S (Z = **15.635**);
LH→LHO (Z = 5.809); LH→T (Z = 12.096)
LHO→LH (Z = 6.682); LHO→T (Z = 4.767); R→L (Z = 4.887); R→LHO (Z = 2.014);
S→Cogn (Z = 3.063); S→L (Z = 2.867); S→R (Z = 3.783); T→LH (Z = 11.322);

5.3 Learning Outcomes Using JLMVR for Japanese Vocabulary and Hiragana

The pre-test and post-test questions for vocabulary and Hiragana in this study were designed based on content from LingoPix and Duolingo. The vocabulary test scores of the participants before and after the experiment are shown in Figure 12. All five participants showed improvement in their scores, with participants 1, 4, and 5 improving by more than 75%. The Hiragana test scores before and after the experiment are shown in Figure 13, where all five participants also showed improvement. Participants 1, 4, and 5 had significant improvements, while participants 2 and 3 only improved from 0 points to 4 and 2 points, respectively, indicating less effective learning.

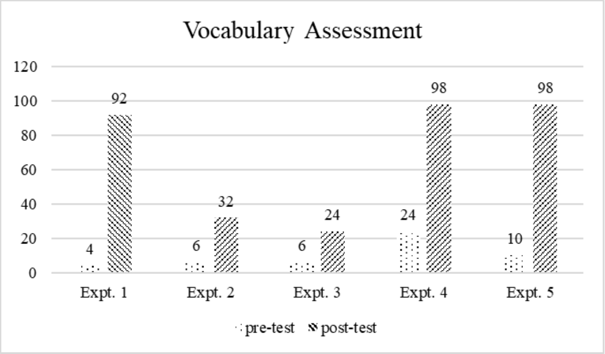

Figure 12: Pre-test and post-test vocabulary scores of participants

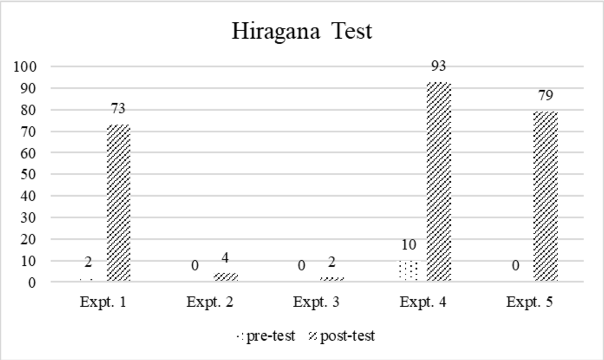

Figure 13: Pre-test and post-test hiragana scores of participants

The behavior patterns diagram of participant 1 in the JLMVR system is shown in Figure 7. The most significant sequence of behaviors was from event T (thinking) to event LH (learn Hiragana) (Z=13.843), with mutual transitions (Z=11.474). Table 2 also shows that the percentage of occurrences for event T (402, 9.3%) and event LH (400, 9.3%) were close, indicating that participant 1 would pause learning when encountering unfamiliar Hiragana and review related knowledge points to reinforce learning impressions. This learning strategy led to an improvement in Hiragana test scores from 2 to 73 points. The next significant transitions were from event L (listening) to event S (speaking) (Z=11.64) and from event S to event R (repetition) (Z=9.894). Events S (1098, 25.4%), L (818, 18.9%), and R (588, 13.6%) were the three most frequent events, indicating that the first transition was for imitation to establish a foundation in the second language, and the second transition was for reinforcing and consolidating memory to convert short-term memory into long-term memory. As a result, participant 1's vocabulary test scores improved from 4 to 92 points. When using Duolingo, participant 1 could quickly identify and click the correct Hiragana pronunciation and remembered the Japanese word for "dog" (いぬ: i nu) encountered in the virtual reality system. However, they made an error due to forgetting the Japanese word for "dirty" (きたない: kitanai) in the Duolingo test, ultimately passing the second level.

The behavior patterns diagram of participant 2 in the JLMVR system is shown in Figure 8. The most significant sequence was from event C (controlling VR) to event L (listening) (Z=12.177). Table 6 shows that the frequencies of event C (1167, 27.0%) and event L (749, 17.3%) were the highest, indicating that participant 2 was often distracted in the system and had to refocus to listen again. Although the frequency of event LH (learn Hiragana) was not low (519, 12%), the adjusted residuals (Z-value) of the transition from event LH to event C was higher than from event LH to other events, suggesting that participant 2 might feel bored or unable to remember when learning Hiragana, leading to wandering or daydreaming. Consequently, the Hiragana test scores only improved from 0 to 4 points. Notably, although the frequency of event R (repetition) was low (216, 5%), the transition from event S (speaking) to event R was also significant (Z=8.934), indicating that during speaking, participant 2 would repeat to consolidate short-term memory, leading to an improvement in vocabulary test scores from 6 to 32 points. When using Duolingo, participant 2 made an error by choosing the answer "よる: yoru" similar to "よむ: yomu." Ultimately, participant 2 passed the first level.

The behavior patterns diagram of participant 3 in the JLMVR system is shown in Figure 10. Although the most common sequence was from event LHO (learn Hiragana with objects) to event T (thinking) (Z=19.424), Table 9 shows that the frequencies of events LHO (170, 3.9%) and T (34, 0.8%) were low, indicating that this learning pattern was not frequently used. Secondly, there was mutual transition between events L (listening) and S (speaking) (Z=14.655 & Z=10.681, respectively). The transition from event C (controlling VR) to event L (Z=13.443) was highly significant, and event C had the second highest occurrence rate (941, 21.8%), indicating that participant 3 often performed operations unrelated to learning and, like participant 2, needed to refocus. Despite needing frequent attention shifts, participant 3 could refocus on learning activities and translate the information heard into speaking behavior. However, the frequency of event R (repetition) was very low (45, 1%), leading to easy forgetfulness of learned vocabulary. Consequently, vocabulary test scores improved from 6 to 24 points, and Hiragana test scores improved from 0 to 2 points. When using Duolingo, participant 3 needed to rely on auxiliary listening to identify Hiragana and made an error in the test for the Hiragana "の: no" without auxiliary listening, ultimately staying at the first level.

The behavior patterns diagram of participant 4 in the JLMVR system is shown in Figure 10. Like participant 1, there was a significant mutual transition between events LH and T (Z=18.641 & Z=16.249, respectively). Additionally, Table 12 shows that the frequencies of events LH and T were also high (743, 17.2%; 371, 8.6%), indicating that participant 4 actively thought about the writing and appearance of related Japanese characters while learning Hiragana, thereby enhancing memory. There was also a mutual transition between events L and S, and the transition from event R to event L was significant, indicating that participant 4 would confirm the correctness of answers through listening after speaking and repeated practice. This learning style led to an improvement in vocabulary test scores from 4 to 98 points and Hiragana test scores from 10 to 93 points. Notably, there was also a mutual transition between events Cogn (cognition) and C, indicating that participant 4 experienced boredom or fatigue near the end of the learning process. When using Duolingo, participant 4 could accurately read and select the corresponding Hiragana symbols for "3,2,1: さん,に,いち" after listening to the prompt, ultimately passing the third level.

In Figure 11, the most significant behavior pattern for participant 5 was the transition from event L to event S. Similar to participants 1 and 4, there was a mutual transition between events LH and T. Unlike other participants, participant 5 had a mutual transition between events LHO and LH. The transition from LHO to LH indicates that participant 5 could deepen the understanding and memory of Hiragana from multiple angles and associate objects with Hiragana. When transitioning from LH to LHO, participant 5 could use the kana chart to recall which objects were composed of which Hiragana, further solidifying mastery of Hiragana symbols. Consequently, vocabulary test scores improved from 10 to 98 points, and Hiragana test scores improved from 0 to 79 points. When using Duolingo, participant 5, like participants 1 and 4, could quickly identify the correct pronunciation of Hiragana but made errors when reading new words, ultimately passing the second level.

In summary, the behavior results of the five participants learning Japanese vocabulary in VR show that participants 1 to 5 exhibited different learning patterns and characteristics in the JLMVR system. Participant 1 would stop to review knowledge points when encountering unfamiliar Hiragana, significantly improving scores. Participant 2 was easily bored or forgetful during learning, often distracted. Participant 3 frequently adjusted the VR equipment, interrupting the learning process. Participant 4's learning strategy included speaking and repeated practice, confirming learning outcomes through listening, demonstrating a positive learning attitude and effective memory methods. Participant 5 deepened understanding and memory of Hiragana through object associations, significantly increasing pre-test and post-test scores. Participants 1, 4, and 5 showed positive thinking and learning behavior patterns in the transitions between learning Hiragana and thinking, achieving high scores in vocabulary tests. In contrast, participants 2 and 3 often needed to refocus, facing distractions during learning, particularly participant 3, whose frequent VR equipment adjustments might have affected learning efficiency. Participant 5 could deepen the understanding of Hiragana from multiple angles, demonstrating a unique learning strategy.

5.4 Flow Scores Using JLMVR

Participant 1 provided positive feedback on the novelty of the learning environment but highlighted the issue of long learning durations. They felt that learning efficiency and interest decreased over extended periods, especially when learning a large amount of Japanese, leading to fatigue and boredom. Thus, they suggested adjusting learning durations based on the intensity of the content, considering four hours sufficient for this experiment to avoid excessive fatigue while maintaining motivation and efficiency. This suggestion emphasizes the need to balance learning time and content volume when designing VR learning environments to optimize learner experience and outcomes.

Participant 2 initially found VR to be an intriguing learning tool but also suggested challenges and improvements for the flow experience in learning Japanese vocabulary and kana in VR. They found VR learning to be a significant challenge for beginners due to multitasking demands, which could distract attention, particularly over extended periods. Participant 2 mentioned that traditional book learning might be more suitable, providing a more direct and focused learning approach. They recommended testing shorter learning sessions in VR to improve efficiency and focus.

Participant 3's flow experience with JLMVR showed mixed feedback. While they initially remembered some basic Hiragana and vocabulary, they struggled with fully understanding phonetics. Long-term use of VR equipment caused discomfort, and the monotonous content reduced learning satisfaction, making the process seem dull. Despite these barriers, they recognized the novelty of learning through direct interaction with objects. However, their overall interest in VR learning was low, lacking motivation and goal-oriented control, making time seem prolonged, explaining the significant transition from event C to event L.

Participant 4 provided both positive and negative feedback on learning Japanese vocabulary in VR. They supported the interactivity and innovation of VR, particularly the simplicity and direct interaction with objects, enhancing learning interest and practicality. However, they noted that the process became monotonous over time, reducing learning efficiency. Participant 4 suggested adding social interaction features for more encouragement and recognition to maintain learning motivation.

Participant 5 had a mixed experience of challenges and achievements in VR learning. They felt accomplished in learning basic Hiragana and vocabulary and noted that time seemed to pass quickly when focused. However, prolonged learning sessions reduced focus and made learning tedious. While they found vocabulary learning easy, Hiragana learning was relatively difficult. Overall, Participant 5 hoped for more interactivity in VR learning, especially in learning Hiragana, suggesting incorporating writing features to enhance efficiency and interest.

In conclusion, this study explored the flow experiences of five participants using JLMVR. The results showed that participants generally responded positively to VR as a novel learning tool, with its enhanced interactivity increasing learning interest. However, they consistently pointed out that long learning durations led to decreased focus and interest, suggesting adjusting learning durations based on content and learner reactions. This finding aligns with flow theory Csikszentmihalyi(2000), where balancing skills and challenges allows entry into a flow state. If challenges are too great, learners may feel anxious, as seen in participants 2 and 3 learning Hiragana. Conversely, if challenges are insufficient, learners may feel bored, as seen in participants 1, 4, and 5 after four hours of learning. Maintaining a balance between skills and challenges is key to achieving flow. Enhancing interactivity and adjusting learning challenges in difficult areas, such as learning Hiragana, are essential to improve learning efficiency and motivation. This study provides insights into the potential and factors of VR in non-Roman script language learning and how to maintain flow states through enhanced interactivity and balanced challenges. Future research can further explore optimizing VR learning environments to meet different learners' needs.

6. Conclusions

This study primarily investigates the effectiveness of applying VR technology in learning Japanese vocabulary and the fifty-syllable alphabet. Through experimental observations of five participants, differences in learning behaviors and performances among the experimenters were analyzed, leading to the following conclusions: Experimenter 1's actively used VR for learning Japanese vocabulary, frequently employing repetitive practice to reinforce learning, with a particular preference for speaking practice and repetition. Experimenter 2 tended to become easily bored or forget the learning content, often displaying distraction behaviors that affected learning efficiency. Experimenter 3's learning process was frequently interrupted due to frequent adjustments of the VR equipment, impacting the continuity and effectiveness of learning. Experimenter 4 adopted a positive learning attitude, preferring to enhance learning outcomes through listening and repetitive practice, demonstrating effective memory techniques. Experimenter 5 strengthened their understanding and memory of the Hiragana by linking objects and Hiragana characters from multiple perspectives, resulting in significant improvements in pre- and post-test scores.

Regarding the flow experience, although all experimenters entered a flow state at some point, potentially due to factors brought by VR, Experimenter 1, Experimenter 4, and Experimenter 5 found that as their skills improved and time progressed, the challenges they faced were no longer sufficient, gradually leading them out of the flow state. In contrast, Experimenter 2 and Experimenter 3 felt that the current materials were too difficult, presenting excessive challenges, and thus were unable to maintain a flow state.

This study explores the impact of integrating VR into learning Japanese vocabulary and the fifty-syllable alphabet, finding that VR can provide a substantial flow experience. However, an appropriate level of challenge is crucial for maintaining this state. Therefore, the difficulty of the materials should be adjusted according to individual abilities to maintain optimal learning outcomes.

Combining the observations and analyses of this study, VR technology has shown its unique advantages in language learning, particularly in enhancing learners' motivation and immersive experiences. The experimenters' experiences indicate that appropriate interactive design and equipment technology can significantly enhance learning efficiency and memory depth. However, individual differences may affect the achievement of flow states. In the future, VR has the potential to become an essential tool in language learning, providing learners with a more efficient and personalized learning experience.

Acknowledgements

This research was supported by the project of teaching practice research program from the Ministry of Education under contract number PGE1110466.

References

[1]. Abich IV, J., Parker, J., Murphy, J. S., & Eudy, M. (2021). A review of the evidence for training effectiveness with virtual reality technology. *Virtual Reality, 25*(4), 919-933. https://doi.org/10.1007/s10055-020-00498-8

[2]. Ader, L. G. M., Crowley, K., Kuhn, S., Caraffini, F., Altındağ, T., & Colreavy-Donnelly, S. (2023). Extended Reality, Augmented Users, and Design Implications for Virtual Learning Environments. *2023 IEEE International Symposium on Technology and Society (ISTAS)*. https://doi.org/10.1109/ISTAS57930.2023.10305947

[3]. Al Farsi, G., Yusof, A. b. M., Romli, A., Tawafak, R. M., Malik, S. I., Jabbar, J., & Bin Rsuli, M. E. (2021). A Review of Virtual Reality Applications in an Educational Domain. *International Journal of Interactive Mobile Technologies*, *15*(22). https://doi.org/10.3991/ijim.v15i22.25003

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

[5]. Azuma, R. T. (1997). A survey of augmented reality. *Presence: teleoperators & virtual environments*, *6*(4), 355-385. https://doi.org/10.1162/pres.1997.6.4.355

[6]. Bakeman, R., & Gottman, J. M. (1997). *Observing interaction: An introduction to sequential analysis*. Cambridge university press. https://doi.org/10.1017/ CBO9780511527685.010

[7]. Cheng, A., Yang, L., & Andersen, E. (2017). Teaching language and culture with a virtual reality game. *Proceedings of the 2017 CHI conference on human factors in computing systems*. https://doi.org/10.1145/3025453.3025857

[8]. Cheng, Y., & Wang, S.-H. (2011). Applying a 3D virtual learning environment to facilitate student's application ability - The case of marketing. *Computers in Human Behavior*, *27*(1), 576-584. https://doi.org/10.1016/j.chb.2010.10.008

[9]. Ciupe, A., Munteanu, S., Parmaxi, A., Nicolaou, A., Gabaudan, O., Nocchi, S., Schalk, A., Orza, B., Campian, C., & Maraciuc, C. (2023). From Speaking Skills to Virtual Mobilities: Challenges of VR Technologies in Communication from the European University of Technology. *2023 IEEE Global Engineering Education Conference (EDUCON)*. https://doi.org/10.1109/EDUCON54358.2023.10125174

[10]. Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational And Psychological Measurement*, *20*(1), 37-46. https://doi.org/10.1177/ 001316446002000104

[11]. Csikszentmihalyi, M. (1991). Flow, the psychology of optimal experience. *American Journal of Psychotherapy*, *45*, 142-143. https://doi.org/10.1176/appi. psychotherapy.1991.45.1.142

[12]. Csikszentmihalyi, M. (1997). Flow and the psychology of discovery and invention. *HarperPerennial, 39*, 1-16.

[13]. Csikszentmihalyi, M. (2000). *Beyond boredom and anxiety*. Jossey-bass. https://doi.org/10.1037/10516-164

[14]. Csikszentmihalyi, M. (2014). *Applications of flow in human development and education*. Springer. https://doi.org/10.1007/978-94-017-9094-9

[15]. Daghestani, L., Al-Nuaim, H., Xu, Z., & Ragab, A. H. M. (2012). Interactive virtual reality cognitive learning model for performance evaluation of math manipulatives. *Journal the King Abdulaziz University Computer and Information Technology Science*, *1*(1), 31-52. https://doi.org/10.4197/Comp.1-1.2

[16]. Dale, E. (1969). Audiovisual methods in teaching. https://eric.ed.gov/?id=ED043234

[17]. Dewey, J. (1916). Nationalizing education. *Journal of Education*, *84*(16), 425-428. https://doi.org/10.1177/002205741608401602

[18]. Erhel, S., & Jamet, E. (2019). Improving instructions in educational computer games: Exploring the relations between goal specificity, flow experience and learning outcomes. *Computers in Human Behavior*, *91*, 106-114. https://doi.org/ 10.1016/j.chb.2018.09.020

[19]. Farshid, M., Paschen, J., Eriksson, T., & Kietzmann, J. (2018). Go boldly!: Explore augmented reality (AR), virtual reality (VR), and mixed reality (MR) for business. *Business horizons*, *61*(5), 657-663. https://doi.org/10.1016/j.bushor. 2018.05.009

[20]. Garcia-Palacios, A., Hoffman, H., Carlin, A., Furness III, T. A., & Botella, C. (2002). Virtual reality in the treatment of spider phobia: a controlled study. *Behaviour research and therapy*, *40*(9), 983-993. https://doi.org/10.1016/S0005-7967(01)00068-7

[21]. Garcia, S., Laesker, D., Caprio, D., Kauer, R., Nguyen, J., & Andujar, M. (2019). An immersive virtual reality experience for learning Spanish. Learning and Collaboration Technologies. *Proceedings of Ubiquitous and Virtual Environments for Learning and Collaboration: 6th International Conference (LCT 2019),Part II 21*. https://doi.org/10.1007/978-3-030-21817-1_12

[22]. Geng, X., & Yamada, M. (2020). The development and evaluation of an augmented reality learning system for Japanese compound verbs using learning analytics. *2020 IEEE International Conference on Teaching, Assessment, and Learning for Engineering (TALE)*. https://doi.org/10.1109/TALE48869.2020.9368345

[23]. Georgiou, Y., & Kyza, E. A. (2018). Relations between student motivation, immersion and learning outcomes in location-based augmented reality settings. *Computers in Human Behavior*, *89*, 173-181.

https://doi.org/10.1016/j.chb.2018.08.011

[24]. Gilakjani, A. P. (2012). The significant role of multimedia in motivating EFL learners' interest in English language learning. *International Journal of Modern Education and Computer Science*, *4*(4), 57. https://doi.org/10.5815/ijmecs.2012.04.08

[25]. Harris, K., & Reid, D. (2005). The influence of virtual reality play on children's motivation. *Canadian Journal of Occupational Therapy*, *72*(1), 21-29. https://doi.org/10.1177/000841740507200107

[26]. Heim, M. (2000). *Virtual realism*. Oxford University Press.

[27]. Hillis, K. (1999). *Digital sensations: Space, identity, and embodiment in virtual reality*. University of Minnesota Press. https://doi.org/10.1177/120633120000300406

[28]. Horwitz, E. K., Horwitz, M. B., & Cope, J. (1986). Foreign language classroom anxiety. *The Modern Language Journal*, *70*(2), 125-132. https://doi.org/10.1177/ 120633120000300406

[29]. Huang, H.-M., & Liaw, S.-S. (2018). An analysis of learners' intentions toward virtual reality learning based on constructivist and technology acceptance approaches. *International Review of Research in Open and Distributed Learning*, *19*(1). https://doi.org/10.19173/irrodl.v19i1.2503

[30]. ITO, H. (2022). Development of Japanese Language Learning Content Using Immersive Virtual Reality. *Inter Faculty*, *11*, 253-260. https://doi.org/10.15068/ 0002003303

[31]. Jia, T., & Liu, Y. (2019). Words in kitchen: An instance of leveraging virtual reality technology to learn vocabulary. *2019 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*. https://doi.org/10.1109/ ISMAR-Adjunct.2019.00-59

[32]. Kiili, I. (2006). Evaluations of an experiential gaming model. *An Interdisciplinary Journal on Humans in ICT Environments*, *2*(2), 187-201. https://doi.org/ 10.17011/HT/URN.2006518

[33]. Kurniawan, B., & Novita, A. (2019). Developing Application Android Based For Japanese Language Learning. *Proceedings of the 1st International Conference on Informatics, Engineering, Science and Technology (NCITEST 2019)*. https://doi.org/10.4108/eai.18-7-2019.2287807

[34]. Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *biometrics*, 159-174. https://doi.org/10.2307/2529310

[35]. Li, C., Liang, W., Quigley, C., Zhao, Y., & Yu, L.-F. (2017). Earthquake safety training through virtual drills. *IEEE Transactions on Visualization and Computer Graphics*, *23*(4), 1275-1284. https://doi.org/10.1109/TVCG.2017.2656958

[36]. Li, X., & Qin, B. (2023). When E-learning Meets Web 3.0: Applications and Challenges. *2023 IEEE 12th International Conference on Educational and Information Technology (ICEIT)*. https://doi.org/10.1109/ICEIT57125.2023.10107827

[37]. Liu, J., & Zou, Y. (2022). An online japanese teaching mode based on virtual reality. *Scientific Programming*, *2022*. https://doi.org/10.1155/2022/2025732

[38]. Liu, M., & Jackson, J. (2008). An exploration of Chinese EFL learners' unwillingness to communicate and foreign language anxiety. *The modern language journal*, *92*(1), 71-86. https://doi.org/10.1111/j.1540-4781.2008.00687.x

[39]. Liu, P., Zhang, Y., & Liu, D. (2022). Flow experience in foreign language writing: Its effect on students' writing process and writing performance. *Frontiers in Psychology*, *13*, 952044. https://doi.org/10.3389/fpsyg.2022.952044

[40]. Liu, R., Wang, L., Koszalka, T. A., & Wan, K. (2022). Effects of immersive virtual reality classrooms on students' academic achievement, motivation and cognitive load in science lessons. *Journal of Computer Assisted Learning*, *38*(5), 1422-1433. https://doi.org/10.1111/jcal.12688

[41]. Marty-Dugas, J., Howes, L., & Smilek, D. (2021). Sustained attention and the experience of flow. *Psychological research*, *85*, 2682-2696. https://doi.org/ 10.1007/s00426-020-01433-x

[42]. Mayer, R. E. (2002). Multimedia learning. *Psychology of learning and motivation*, *41*, 85-139. https://doi.org/10.1016/S0079-7421(02)80005-6

[43]. Oxford, R., Park-Oh, Y., It, S., & Sumrall, M. (1993). Japanese by satellite: Effects of motivation, language learning styles and strategies, gender, course level, and previous language learning experience on Japanese language achievement. *Foreign Language Annals*, *26*(3), 359-371. https://doi.org/10.1111/j.1944-9720.1993.tb02292.x

[44]. Parmaxi, A. (2023). Virtual reality in language learning: A systematic review and implications for research and practice. *Interactive Learning Environments*, *31*(1), 172-184. https://doi.org/10.1080/10494820.2020.1765392

[45]. Peixoto, B., Bessa, L. C. P., Gonçalves, G., Bessa, M., & Melo, M. (2023). Teaching EFL With Immersive Virtual Reality Technologies: A Comparison With the Conventional Listening Method. *IEEE Access*, *11*, 21498-21507. https://doi.org/10.1109/ACCESS.2023.3249578

[46]. Penpinun, C., Al Ahsan, R., Phattana, N., Srisupakwong, K., Noisri, S., Sinpan, N., Mahakornpetch, T.,

Jearavongtakul, S., Sasithong, P., & Phyo, W. (2021). Development of Chinese vocabulary learning by utilizing virtual reality. *2021 36th International Technical Conference on Circuits/Systems, Computers and Communications (ITC-CSCC)*. https://doi.org/10.1109/ITC-CSCC52171.2021. 9501455

[47]. Qian, Y., Prajapati, S., Schwartz, A., Jung, A., Seitz, U., Van Alfen, J., Lewis, L., Kim, M., Kramer, A., & Chukoskie, L. (2023). Integrated aerobic exercise into adult second language learning in virtual reality game. *2023 IEEE Conference on Games (CoG)*. https://doi.org/10.1109/CoG57401.2023.10333177

[48]. Rajesh, R., Chinthamu, N., Rani, S., Kumar, M., & Sivaiah, B. V. (2023). Development of powered chatbots for natural language interaction in metaverse using deep learning with optimization techniques. *2023 Second International Conference on Augmented Intelligence and Sustainable Systems (ICAISS)*. https://doi.org/10.1109/ICAISS58487.2023.10250650

[49]. Richard, T., & Rajakumari, R. (2023). The impact of virtual reality and artificial intelligence on a self-guided online course in oral english for undergraduates. *2023 International Conference on Advances in Computing, Communication and Applied Informatics (ACCAI)*. https://doi.org/10.1109/ACCAI58221.2023.10200338

[50]. Sackett, G. P.A (1974). Nonparametric lag sequential analysis for studying dependency among responses in observational scoring systems. Unpublished manuscript, University of Washington.

[51]. Sarkady, D., Neuburger, L., Egger, R. (2021). Virtual Reality as a Travel Substitution Tool During COVID-19. *In: Wörndl, W., Koo, C., Stienmetz, J.L. (eds) Information and Communication Technologies in Tourism 2021*. Springer, Cham. https://doi.org/10.1007/978-3-030-65785-7_44

[52]. Sherman, W. R., & Craig, A. B. (2003). Understanding virtual reality. *San Francisco, CA: Morgan Kauffman*.

[53]. Soga, M., & Kimoto, K. (2022). A Word Learning Environment with Experiences in VR World. *Procedia Computer Science*, *207*, 3340-3347. https://doi.org/ 10.1016/j.procs.2022.09.392

[54]. Su, Y.-S., Cheng, H.-W., & Lai, C.-F. (2022). Study of virtual reality immersive technology enhanced mathematics geometry learning. *Frontiers in Psychology*, *13*, 760418. https://doi.org/10.3389/fpsyg.2022.760418

[55]. Suh, A., & Prophet, J. (2018). The state of immersive technology research: A literature analysis. *Computers in Human Behavior*, *86*, 77-90. https://doi.org/10.1016/j.chb.2018.04.019

[56]. Suhendi, A. (2018). Constructivist learning theory: The contribution to foreign language learning and teaching. *KnE Social Sciences*, 87-95. https://doi.org/10.18502/kss.v3i4.1921

[57]. Sun, Z., Lin, C.-H., Lv, K., & Song, J. (2021). Knowledge-construction behaviors in a mobile learning environment: A lag-sequential analysis of group differences. *Educational Technology Research and Development*, *69*, 533-551. https://doi.org/10.1007/s11423-021-09938-x

[58]. Sutherland, I. E. (1965). The ultimate display. *Proceedings of the IFIP Congress*.

[59]. Sutherland, I. E. (1968). A head-mounted three dimensional display. *Proceedings of the December 9-11, 1968, Fall Joint Computer Conference, Part I.* 757-764. https://doi.org/10.1145/1476589.1476686

[60]. Symonenko, S., Zaitseva, N., Osadchyi, V., Osadcha, K., & Shmeltser, E. (2020). Virtual reality in foreign language training at higher educational institutions. https://doi.org/10.31812/123456789/3759

[61]. Tai, T.-Y., & Chen, H. H.-J. (2021). The impact of immersive virtual reality on EFL learners' listening comprehension. *Journal of Educational Computing Research*, *59*(7), 1272-1293. https://doi.org/10.1177/0735633121994291

[62]. Tai, T.-Y., Chen, H. H.-J., & Todd, G. (2022). The impact of a virtual reality app on adolescent EFL learners' vocabulary learning. *Computer Assisted Language Learning*, *35*(4), 892-917. https://doi.org/10.1080/09588221.2020.1752735

[63]. Teng, M. F. (2023). The effectiveness of multimedia input on vocabulary learning and retention. *Innovation in Language Learning and Teaching*, *17*(3), 738-754. https://doi.org/10.1080/17501229.2022.2131791

[64]. Tlili, A., Wang, H., Gao, B., Shi, Y., Zhiying, N., Looi, C.-K., & Huang, R. (2023). Impact of cultural diversity on students' learning behavioral patterns in open and online courses: A lag sequential analysis approach. *Interactive Learning Environments*, *31*(6), 3951-3970. https://doi.org/10.1080/10494820.2021. 1946565

[65]. Udjaja, Y., Tanuwijaya, K., & Wairooy, I. K. (2019). The use of role playing game for Japanese language learning. *Procedia Computer Science*, *157*, 298-305. https://doi.org/10.1016/j.procs.2019.08.170

[66]. Vázquez, C., Xia, L., Aikawa, T., & Maes, P. (2018). Words in motion: Kinesthetic language learning in virtual reality. *2018 IEEE 18th International Conference on advanced learning technologies (ICALT)*. https://doi.org/10.1109/ICALT. 2018.00069

[67]. Villena-Taranilla, R., Tirado-Olivares, S., Cozar-Gutierrez, R., & González-Calero, J. A. (2022). Effects of virtual reality on learning outcomes in K-6 education: A meta-analysis. *Educational Research*

Review,*35*, 100434. https://doi.org/10.1016/j.edurev.2022.100434

[68]. Ye, X. (2022). Optimization of Multimodal Japanese Teaching Model Using Virtual Reality. *Mobile Information Systems*, *2022*. https://doi.org/10.1155/2022/7364279

[69]. Yoh, M.-S. (2001). The reality of virtual reality. Proceedings seventh international conference on virtual systems and multimedia. https://doi.org/10.1109/VSMM. 2001.969726

[70]. Yu, D. (2023). AI-Empowered Metaverse Learning Simulation Technology Application. *2023 International Conference on Intelligent Metaverse Technologies &Applications (iMETA)*. https://doi.org/10.1109/iMETA59369.2023.10294830

[71]. Zarzour, H., Bendjaballah, S., & Harirche, H. (2020). Exploring the behavioral patterns of students learning with a Facebook-based e-book approach. *Computers & Education*, *156*, 103957. https://doi.org/10.1016/j.compedu.2020.103957

[72]. Zheng, Y. (2008). Anxiety and second/foreign language learning revisited. *Canadian Journal for New Scholars in Education/Revue canadienne des jeunes chercheures et chercheurs en education*, *1*(1), 1-12.