



Castledown

 OPEN ACCESS

Technology in Language Teaching & Learning

ISSN 2652-1687

<https://www.castledown.com/journals/tltl/>

Technology in Language Teaching & Learning, 7(3), 102540 (2025)

<https://doi.org/10.29140/tltl.v7n3.102540>

Development of Task-Based Language Teaching in Virtual Reality: A Case Study



KOYO OGAWA 

Graduate School of Language & Literature, Prefectural University of Kumamoto, Japan

ogako.1124@gmail.com

Abstract

Virtual reality (VR) has attracted increasing interest for its potential to enhance second language acquisition through immersive and interactive experiences. However, effectively integrating VR with established language teaching pedagogy remains a critical challenge. This study addresses that gap by embedding the principles of task-based language teaching (TBLT) into a custom-designed VR learning environment. Using “making orders in a bar” as a representative target task, the study illustrates the design of pedagogical tasks and the development of an original VR environment using open-source tools. Employing a small-scale case study design, the research investigates whether this VR-based TBLT (VR-TBLT) facilitates learners’ acquisition of the target discourse and supports the transfer of skills to real-world contexts. It also explores learners’ perceptions of the VR-TBLT experience. The findings indicate that the VR-TBLT approach promotes the use of the target language, supports near transfer to real-world tasks through contextual similarity, and enhances learners’ engagement. Participants reported high levels of enjoyment and presence within the VR environment, although minor technological issues were noted. These results underscore VR-TBLT’s potential as a pedagogically grounded tool for language learning and highlight key considerations for future research on scalability and long-term outcomes.

Keywords: virtual reality (VR), VR affordances, transfer of learning, task-based language teaching (TBLT), sense of presence, immersion

Introduction

Recent years have seen growing interest in using virtual reality (VR) for language learning (e.g., Dhimolea et al., 2022; Parmaxi, 2020; Taguchi & Zhao, 2025). A notable trend in this field is the shift

Copyright: © 2025 Koyo Ogawa. This is an open access article distributed under the terms of the Creative Commons Attribution Non-Commercial 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. **Data Availability Statement:** All relevant data are within this paper.

from low-immersion VR, such as desktop-based simulations, to high-immersion VR, which uses head-mounted displays (HMDs) to offer a more immersive and interactive experience (Kaplan-Rakowski & Gruber, 2019). This study focuses on high-immersion VR, henceforth referred to simply as “VR.”

As consumer interest in VR has grown since the release of more affordable HMDs around 2015 and 2016, the application of this immersive technology in language learning began to attract researchers’ attention around 2020 (Taguchi & Zhao, 2025). VR offers a range of affordances that are not available in traditional teaching methods. Frequently cited affordances include interactivity, a high degree of immersion, and a strong sense of presence (e.g., Ciekanski et al., 2020; Taguchi & Zhao, 2025; Yan & Lowell, 2024; Yuditseva, 2023). Interactivity refers to users’ ability to interact with and influence virtual environments. Immersion is made possible by HMDs that block out the real world, creating a seamless virtual experience. This, in turn, generates a sense of presence, often described as the feeling of “being there,” which enables the users to feel physically situated within the virtual environment. Thanks to these distinctive affordances, VR has attracted increasing interest among language educators and researchers as a promising tool for promoting authentic language learning.

However, reviews of VR-based language learning research often report inconclusive findings regarding its effectiveness (e.g., Dhimolea et al., 2022; Parmaxi, 2020; Taguchi & Zhao, 2025; Yuditseva, 2023). While qualitative studies often highlight learner engagement, quantitative studies report inconsistent outcomes (Dhimolea et al., 2022). Echoing these findings, Yuditseva (2023) argues that “Virtual Reality, on its own, may not result in positive outcomes” (p. 9), emphasizing that the effectiveness of VR depends on the implementation of appropriate strategies. A common concern in recent studies is the lack of a robust pedagogical framework guiding the use of VR in language education (Smith & McCurrach, 2021; Taguchi & Zhao, 2025; Yan & Lowell, 2024).

To address this gap, the present study integrates task-based language teaching (TBLT) with immersive VR to create a pedagogically grounded learning environment. It demonstrates how TBLT principles can guide VR content design and evaluates the approach through a small-scale case study. Using both quantitative and qualitative methods, including test comparisons, rubric-based performance evaluations, conversation analysis, and In Vivo coding, the study explores the potential of this VR-TBLT model to support language development and transfer to real-world tasks. The next section reviews how previous studies have addressed language learning tasks in VR and highlights key challenges of integrating VR into language learning contexts.

Challenges of Integrating VR in Language Learning

Oda (2024) examined Japanese university teachers’ perceptions of using VR in language learning through the lens of the Technological Pedagogical Content Knowledge framework. His findings highlight the perceived challenges across three domains of knowledge: technological knowledge, referring to difficulties with operating VR hardware and software; technological content knowledge, relating to the limited availability of accessible and suitable learning content; and technological pedagogical knowledge, which encompasses the need for new teaching approaches that are compatible with VR environments.

Among these challenges, concerns regarding technological content knowledge, particularly the lack of VR content designed for language learning, are frequently emphasized in the literature (e.g., Frazier et al., 2018; Yamanaka et al., 2023; Yano, 2022). A relevant issue is the inadequate quality of the existing language learning software. For instance, Kawasumi and Ishii (2023) investigated learners’ perceptions of MondlyVR, one of the most widely used VR applications for language learning. They reported several limitations, including the inflexibility of the non-player character system in handling unexpected

user utterances and the low accuracy of the speech recognition feature. In response, researchers have recommended the development, design, and customization of original VR learning software and content tailored specifically for language education (e.g., Yamanaka et al., 2023; Yano, 2022).

In addition to technological content challenges, issues concerning technological pedagogical knowledge are also prominent in prior studies. In VR-TBLT studies, researchers (e.g., Hoang et al., 2023; Taguchi & Zhao, 2025) have highlighted the limited number of studies that design or discuss VR-based language teaching within a solid TBLT framework. As Taguchi and Zhao (2025) noted, “of the technological tools examined in the TBLT literature, VR has received the least attention” (p. 26). Similarly, in proposing principles for VR-based TBLT, Yan and Lowell (2024) observed that “research has highlighted a lack of consideration for pedagogy and virtual reality affordances in designing learning activities within VR environments” (p. 309). Although many studies claim to use task-like activities in VR language instruction, they often lack a clearly defined pedagogical foundation or fail to articulate it explicitly.

In summary, while existing research has thoroughly identified the technological, content, and pedagogical challenges in VR-based language education, few studies have provided practical, theoretically grounded examples of how to overcome them. In particular, the lack of custom-made, pedagogically sound VR learning content remains a significant gap. This study addresses these issues by designing an original VR environment guided by established TBLT criteria (Taguchi & Zhao, 2025). To do so, the next section discusses previous research focusing more on the integration of VR and TBLT.

Integration of TBLT and VR: Potential and Challenges

Although theoretically and pedagogically grounded VR content for language learning remains scarce, recent research has highlighted the promising potential of integrating VR with TBLT. This is largely due to the unique affordances of VR that align well with TBLT principles (Hoang et al., 2023). For example, Taguchi and Zhao (2025) suggested that immersive features of VR, such as presence, embodiment, and multimodal input, can help tasks become more meaning-focused, contextualized, and authentic. Similarly, Yan and Lowell (2024) argued that the learner-centered nature of TBLT matches well with VR’s affordances, such as authenticity and interaction. Christoforou et al. (2024) also emphasized that both TBLT and VR can bring authenticity into language education.

TBLT is an experiential, learning-by-doing approach that emphasizes the use of language for real-world communication (Khezlou, 2025; Taguchi & Zhao, 2025). It promotes language acquisition through engagement in meaningful tasks that resemble real-life situations. The core tenets of TBLT include a focus on meaning over isolated linguistic forms, authenticity, and the completion of non-linguistic outcomes—goals beyond simply demonstrating language knowledge (Taguchi & Zhao, 2025).

VR provides an ideal environment to enact these principles. The immersion and presence in VR can create highly contextualized, authentic settings for communication (Christoforou et al., 2024; Taguchi & Zhao, 2025), allowing learners to feel as though they are genuinely “there.” Furthermore, VR platforms support interaction and collaboration—key elements in TBLT—by enabling peer communication and joint problem-solving in shared virtual spaces (Smith & McCurrach, 2021; Wiboolyasarin et al., 2023; Taguchi & Zhao, 2025). These interactions foster creativity, discovery, and meaning negotiation, all of which are vital to task-based learning.

The kinesthetic and interactive nature of VR also enhances experiential learning. Learners can explore virtual environments and manipulate objects to complete tasks, applying the target language in real-

istic contexts (Smith & McCurrach, 2021; Taguchi & Zhao, 2025). Such embodied experiences help solidify the connection between language and its practical use.

Despite this synergy, the integration of VR into TBLT is still in its early stages. Khezrlou (2025) argued that technology-mediated TBLT needs more research on theoretical issues like task features and their implementation. Ondarra et al. (2025) also posited that teachers should be aware of the specific affordances of the technology and consider appropriate pedagogy when implementing it and designing tasks. A key limitation noted in prior research is that many VR-based language tasks fall short of realizing one of TBLT's defining characteristics: achieving non-linguistic outcomes (Taguchi & Zhao, 2025). Although many tasks in VR environments are meaning-focused and authentic, they often prioritize the practice of specific language forms rather than requiring learners to complete a real-world goal.

This shortfall is not due to technological limitations but to shortcomings in pedagogical design. Researchers and educators must go beyond simply leveraging VR's affordances and intentionally design tasks that fully reflect TBLT principles. This includes ensuring that the tasks are goal-oriented, contextualized, and focused on real-world, non-linguistic outcomes.

In sum, while VR holds a significant promise for enriching language education through its immersive, interactive, and experiential qualities, realizing its full potential within TBLT frameworks depends on thoughtful, theory-informed task design. Practitioners must take an active role in aligning VR-based tasks with the foundational principles of TBLT to create truly effective and meaningful learning experiences. The current study directly addresses this issue through a pioneering endeavor demonstrating the integration of all the key elements of TBLT with a custom-designed VR environment.

The Current Study

This study comprises two main components: designing a model VR-TBLT and an exploratory case study. The first component aims to advance VR-TBLT research by addressing the gaps identified in the existing literature, explicitly showing how the TBLT can be integrated into VR environments. By illustrating the process of designing both a task-specific VR environment and the corresponding pedagogical tasks, this study sought to encourage second language educators to develop their own VR-TBLT systems tailored to the needs of their own learners and instructional settings.

The second component is an exploratory case study that investigates the potential of the custom-designed VR-TBLT system to enhance learners' task-related language skills. However, this study is not a controlled experiment and does not include a comparison group, such as learners exposed to non-VR TBLT. As such, the unique contribution of VR cannot be empirically isolated, and the findings should be interpreted with this limitation in mind. Furthermore, the small number of participants, namely two adult males with varying levels of English proficiency, precludes statistical generalization. Instead, the focus is on detailed qualitative observations of the learners' language use within the specific VR-TBLT context and their ability to transfer language skills to a real-world task.

Despite these limitations, the case study offers rich insights into how a custom-built VR-TBLT system can support language development in specific contexts. The goal is not to make broad claims about VR-TBLT's effectiveness but to provide a well-documented example that other educators and researchers may find useful when designing or adapting their own VR-based language learning environments.

To investigate the potential impact of the VR-TBLT system, this study includes a pretest, a post-test, and a real-world transfer test, focusing on both language development and task performance. This dual focus responds to concerns raised in the literature that technology-mediated TBLT studies

often prioritize linguistic gains at the expense of task performance quality (Kim & Namkung, 2024). Additionally, the learners' perceptions of their experience with VR-TBLT were gathered through a post-treatment questionnaire and analyzed qualitatively.

The case study addresses the following research questions:

1. To what extent can the custom VR-TBLT contribute to the acquisition of task-related language skills?
2. How do learners perceive their English learning experiences in VR-TBLT?

Methodology

Designing the VR-TBLT system

For second language learners, common real-world communicative needs include tasks such as shopping, eating out, and using public transportation (Iizuka, 2019). This study focuses on the target task of making orders in a bar and designed six pedagogical tasks (PT1–PT6) to support its development. Pedagogical tasks (PTs) are structured classroom activities that help learners acquire the language skills necessary to perform real-life target tasks (Nunan, 2010). Examples of PTs include jigsaw tasks, information-gap activities, problem-solving tasks, decision-making, and exchanging opinions (Jackson, 2022).

Although definitions of a “task” vary among TBLT researchers, Taguchi and Zhao (2025) identified three widely accepted characteristics: tasks should be (1) meaning-oriented, (2) grounded in authentic real-life contexts, and (3) should produce non-linguistic outcomes. However, after reviewing 33 VR-TBLT studies published over the past decade, Taguchi and Zhao (2025) reported that many of the VR tasks used in these studies fail to meet these three core criteria. Of the 39 VR tasks analyzed, only 12 met all of the requirements. It is therefore essential to carefully incorporate these principles when designing TBLT for VR environments.

The tasks in this study fulfill the three criteria as follows. First, they are meaning-oriented: learners use language as a communicative tool to accomplish the given tasks rather than explicitly studying grammar or vocabulary. Target discourse items are introduced implicitly through 360° modeling videos that learners watch before each task. In addition, useful expressions are listed on task handouts, allowing the learners to refer to them when needed to express their thoughts or complete the task. Second, the tasks are authentic. Both the context and the discourse are drawn from real-life ordering interactions in an actual bar. The VR environment in this study is a faithful reproduction of a real-world bar, and the target discourse was collected from native English speakers placing orders in that bar. Third, the tasks produce non-linguistic outcomes. These include insights into the US ordering system, which differs from that of the learners' home country; understanding tipping culture, which is absent in their native context; practicing decision-making that involves calculating in a foreign currency; and achieving tangible results, such as receiving drinks and food in a bar.

The next section outlines the development of the VR learning environment, followed by detailed descriptions of the pedagogical tasks (PT1–PT6) used in this study.

Building a VR Environment for TBLT

To maximize authenticity and relevance to real-life tasks, the VR learning environment used in this study was designed as a digital replica of an actual bar. This reproduction began with the use of a

LiDAR scanner (available on iPhone Pro and iPad Pro devices) to capture three-dimensional (3D) data of the physical space. Several mobile apps make it possible to scan and download 3D objects at no cost, making this process accessible to educators who wish to recreate specific real-world locations, such as a restaurant, hotel counter, or university store.

The next step involved building the VR environment using 3D modeling software. This study used Blender (<https://www.blender.org/>), a widely used, free, open-source modeling tool. The scanned data were imported into Blender as 3D object files (e.g., OBJ or FBX), allowing the digital bar to be modeled with the exact dimensions and layout of the real one. Key elements, such as furniture, drinks, and utensils, were individually modeled to replicate the physical bar in as much detail as possible (Figure 1). This modeling process took approximately two months, during which the author self-taught himself Blender using beginner-friendly tutorials available on YouTube. These resources make it possible for anyone to begin modeling their own VR learning environments with minimal prior experience.

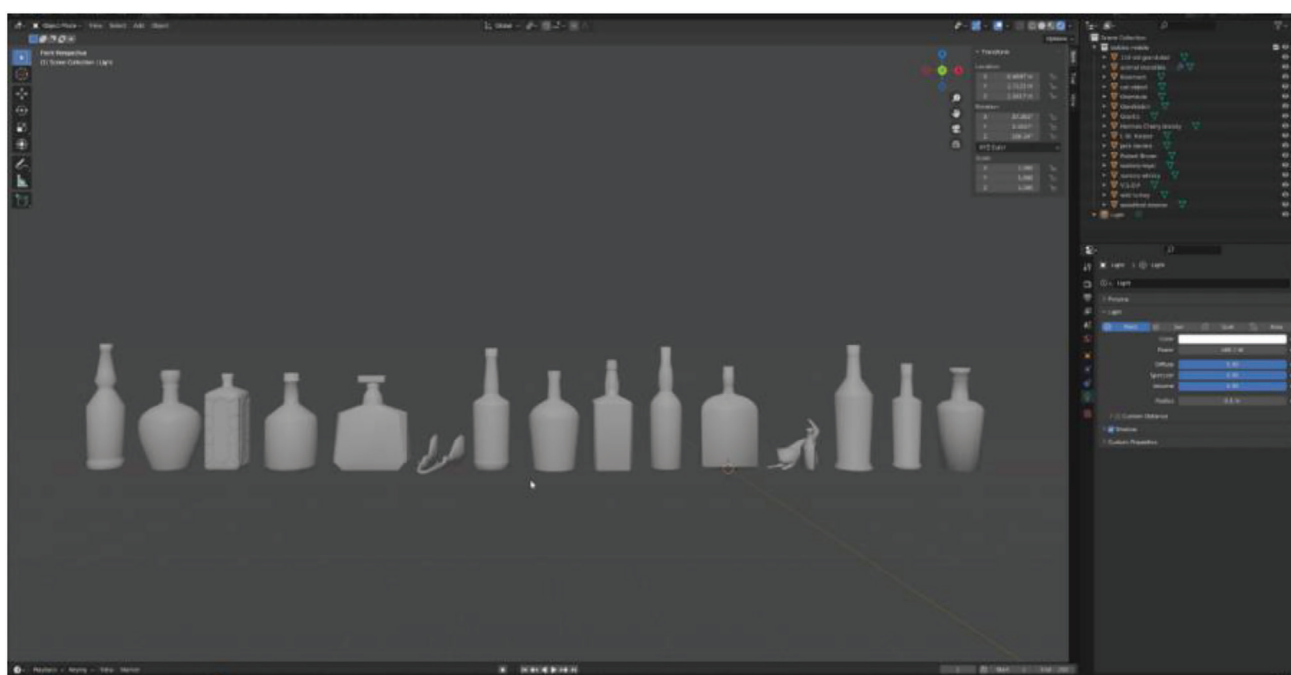


Figure 1 Screenshot of Modeling in Blender.

Once the modeling was complete, the bar and the associated 3D objects were exported to Unity (<https://unity.com/ja>), a popular game engine often used for VR and game development. Unity was used because VRChat, the Metaverse platform employed here, supports only specific versions of Unity. Within Unity, the 3D environment was enhanced with features such as realistic shadows and object collision, enabling interactive functions like picking up virtual bottles or glasses. After converting 3D objects into interactable game assets in Unity, the VR bar was uploaded to VRChat, one of the most accessible and user-friendly Metaverse platforms available. VRChat supports custom avatars and environments, offering significant flexibility for educational use. Numerous tutorial videos on YouTube guide users through the process of making 3D data playable and uploading it to VRChat.

The final environment is accessible through both HMDs and standard desktop or laptop computers. When using HMDs and controllers, learners can move freely within the space, look around naturally, and interact physically with virtual objects, such as picking up glasses or carrying drinks (Figure 2). These features enable a high level of interactivity: one of VR's key affordances. The following section

describes the pedagogical tasks designed to take place in the VR bar and shows how each embodies the essential principles of TBLT.

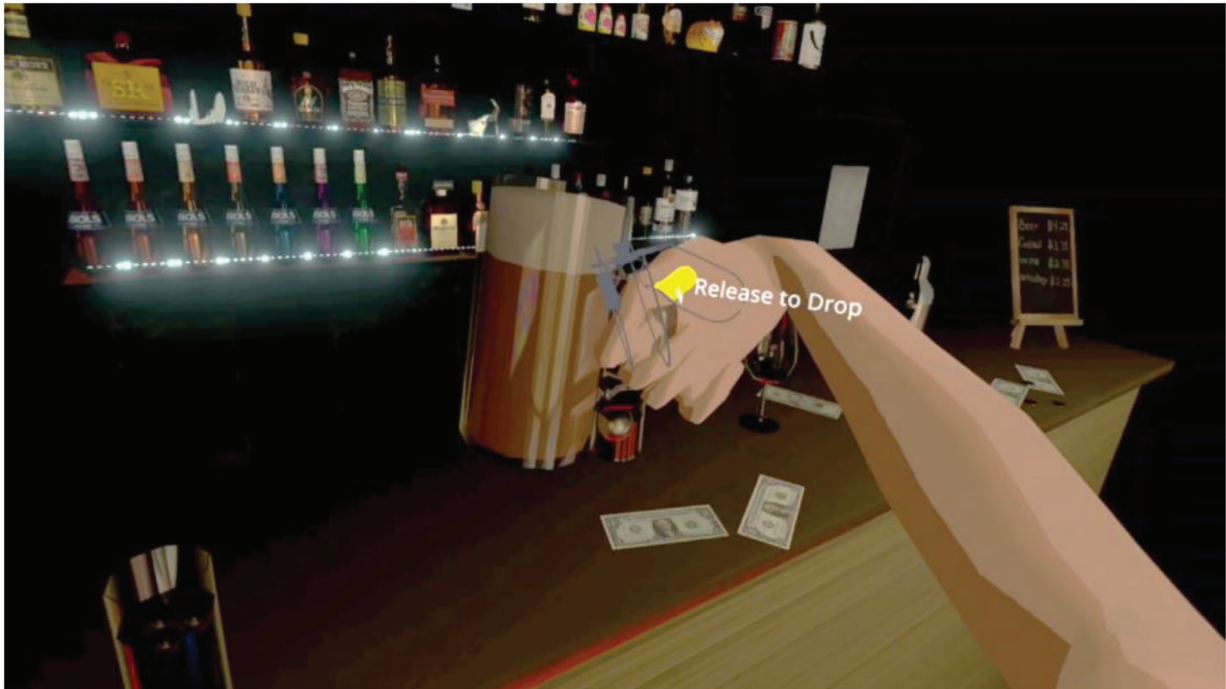


Figure 2 *Interacting with Virtual Objects in VRChat.*

Pedagogical Tasks

TBLT in this study consists of six PTs including input tasks, interactive tasks, and role-play simulations. All the tasks require VR headsets and take about four hours to complete (Figure 3).



Figure 3 *A Participant Wearing a VR Headset.*

PT1: Learners complete a fill-in-the-blank worksheet while watching a 360° VR video using VR headsets that show native speakers ordering at a bar. This task introduces cultural differences, such as tipping and ID checks, and aligns with the recommendations of Ellis (2003) and Shintani (2018) to start with low-complexity tasks. The video, uploaded with subtitles, allows learners to pause, refer to subtitles, and replay as needed (see Appendix A for the handout). PT1 takes about 20 minutes to complete.

PT2: Participants use VR headsets to watch a 360° video explaining US currency and complete a worksheet identifying the figures and values on bills and coins (Appendix B). They then enter a virtual bar to practice identifying the correct payment for menu items. This collaborative task lets learners interact in English while managing virtual US money (Figure 4). PT2 takes about 30 minutes to complete.



Figure 4 Participants Trying a PT in VR.

PT3: After viewing a 360° model video via VR headsets that show a customer discussing alcohol content with a bartender, the participants complete an information-gap task in VR. Using distinct handouts with missing details (Appendix C), learners alternate roles as a customer and the bartender, exchanging information to complete their forms. PT3 takes about 30 minutes.

PT4: Participants wear headsets and watch a 360° video where a native speaker models how to tip and learn about the tipping amount (typically 20% of the bill). A worksheet (Appendix D) helps learners decide on tips for drinks, after which they role-play in VR to share and compare tipping amounts. This activity combines the exchange of opinion and practical application. PT4 takes about 30 minutes.

PT5: Focused on cocktails, this task involves a 360° video showing a customer asking about a drink's ingredients and taste. After watching the video using VR headsets, participants complete an information-gap activity in VR which let them role-play as the bartender and a customer to exchange details about various cocktails and fill out a worksheet (Appendix E). PT5 takes about 40 minutes.

PT6: The final task is a comprehensive role-playing simulation. After watching a 360° model video using VR headsets, the participants independently order drinks and appetizers in a virtual bar while adhering to constraints like budget limits and avoiding duplicate orders. They also consider a friend's preferences. This open-ended task consolidates the skills developed in earlier PTs and mirrors the pre-test with slight modifications in budget and preferences. PT6 takes about an hour.

Each task progressively builds language skills, starting from input tasks (PT1 and PT2) then proceeding to interactive tasks (PT3–PT5), and culminates in a complete role-play simulation (PT6). The following sections describe a small case study that tested these PTs and the VR bar to see to what extent the custom-made VR-TBLT can facilitate participants' language development as well as their transfer to real-world contexts. It also sheds light on the participants' perceptions of their learning experiences in the VR-TBLT tasks.

Participants

Two Japanese-speaking male participants, both 26 years old, took part in this study. Participant A has limited English exposure, having only studied it in school without any overseas experience. His estimated Common European Framework of Reference for Languages (CEFR) level is A1–A2, according to a free online proficiency test, the EF Standard English Test (EF SET), developed by English First (<https://www.efset.org/ja/>). He lacks confidence in English as a result of minimal effort in school but shows interest in learning for potential travel opportunities.

With 17 years of English study, Participant B spends over 4 hours daily preparing for an upcoming study-abroad program in the US. His estimated CEFR level is B1–B2 according to the online proficiency test. Although he has some international experience, including visits to English-speaking countries, he describes his conversational confidence as moderate but sees learning English as essential.

Materials

The study employed PICO 4 VR headsets (<https://www.picoxr.com/global/products/pico4>) to provide the participants with a fully immersive English language learning experience. These headsets support six degrees of freedom, allowing users to move their entire body, not just their heads, within the virtual environment. At the time of the study, the PICO 4 was one of the most affordable standalone headsets offering six degrees of freedom. It also outperformed other comparable VR devices, such as the Meta Quest 2, in terms of resolution, field of view, and weight balance.

The 360° videos used in all PTs were recorded using the Insta360 X3 camera, which captures every angle of a scene. This allows users to look freely in all directions while watching, creating a strong sense of presence and immersion. These videos were made available through YouTube and can be viewed via the VR headset with both English and Japanese subtitles.

In addition to the videos, several handouts were created and distributed to participants to support their engagement with the PTs (see the Appendices).

Procedure

The study involved three phases: pre-task, treatment, and post-task (see Figure 5). The steps of the procedure were carried out over four separate days: one day for the pre-task phase, two days for the treatments, and one day for the post-task phase.



Figure 5 *Procedure.*

Pre-Task Phase

Participants complete two pretests: an expression test and a task performance test.

A) Expression Test

This test measures the participants' input and output language skills via self-assessment.

Input assessment: Participants watch seven videos of native speakers using English expressions and rate their comprehension on a 4-point scale (1 = no understanding; 4 = full understanding). These test videos were created by extracting relevant scenes from longer recordings in which native speakers perform the task. The full-length versions were used in the PTs as modeling videos.

Output assessment: Participants see Japanese sentences and rate how well they believe they can express these sentences in English using the same scale (1 = no idea; 4 = confident). Subjective evaluation prevents unintentional learning from direct translation, minimizing the practice effect before the treatment.

B) Task Performance Test

The participants perform a role-play task in VR, simulating real-world interactions.

Task: Ordering food and drinks at a virtual bar within a budget (see Appendix F for the task's instructions).

Setup: Participants wear HMDs and act as customers to complete the task in VR while the researcher acts as the bartender using a laptop from a physically different room from the participant in the same building. Performances are screen-recorded and transcribed for later comparison with the post-treatment results. To avoid practice effects, no real-world role-playing test occurs at this stage.

Treatment Phase

The participants engage in PT1 to PT6, which were designed earlier. All PTs require HMDs to complete the given tasks, and most parts of PT2–PT6 take place in the VR bar created earlier. To mitigate the participants' fatigue from using HMDs for a long time, the tasks are divided into two sections (PTs1–3 and PTs4–6) and conducted on two separate days.

Post-Task Phase

The participants complete an expression post-test, a real-world task (RWT), and an exit questionnaire to evaluate their learning outcomes and perceptions.

(A) Expression Post-Test

This test replicates the pretest format to measure improvements in participants' input and output knowledge. Since feedback was withheld during the pretest, the results directly reflect the treatment's impact.

(B) Real-World Task

This transfer test assesses the participants' ability to apply their learning in real-world contexts. Participants perform the same role-play task as in the pretest but in a physical bar in the real world (Figure 6).

Task: The participants use real US currency to order food and drinks within a budget (see Appendix F for the task's instructions).

Setup: An international student, asked to speak English at a normal speed, plays the role of the bartender, and a Japanese male volunteer accompanies the participants as a non-English-speaking friend. Unlike the role-play in the pretest, everything was physical and real in this test: The bartender serves real drinks and food while the participants pay real US money. The performances are video-recorded for transcription and comparison with virtual bar interactions.



Figure 6 *RWT in a Real Bar.*

(C) Exit Questionnaire

The participants complete a reflective questionnaire (in Japanese) to provide feedback on their VR-TBLT experience. The questions are listed in Appendix G.

Analysis

To address the first research question (To what extent can the custom VR-TBLT contribute to the acquisition of task-related language skills?), this study focused on two key aspects: (1) language gains achieved through VR-TBLT and (2) the transfer of acquired skills to a RWT.

Language development in this study was primarily examined through pre- and post-treatment expression tests based on learners' self-assessments. In these tests, participants rate their ability to understand

and express the given target expressions in English on a Likert-type scale. These expressions correspond to the target discourse integrated into the PTs, allowing the test to directly assess the language intended for acquisition. Changes in self-assessment scores are therefore used as an indicator of the participants' progress in acquiring task-relevant language skills.

While self-assessment is inherently subjective and may be prone to bias, Kostopoulou (2009) points out that a growing body of research has confirmed the validity and reliability of learner self-assessments by comparing them with external evaluations such as teacher assessments. To further strengthen the credibility of the findings, this study also incorporated conversation analysis. Specifically, the participants' spoken output in the final task (PT6) was compared with their performance in the pre-task phase, enabling the identification of actual changes in language use. By combining self-assessment with objective analysis of language use, the study aimed to provide a more balanced picture of learners' language development.

Skill transfer to a RWT was evaluated by four raters (the author and three experienced English educators) using a task performance rubric developed for this study (see Appendix H). To move beyond binary completion metrics (e.g., Nolen, 2020), the rubric uses a 10-point scale to allow for more nuanced judgments of communicative effectiveness. Raters assessed the participants' performance in both PT6 via screen recordings and the RWT via video recordings.

The initial inter-rater reliability, calculated using the intraclass correlation coefficient (ICC), was moderate (0.545), which raises concerns about consistency. To improve reliability for the present analysis, the ratings from the three most closely aligned raters were retained, excluding one rater as an outlier. This adjustment increased the ICC to 0.683, a level interpreted as "good" (Cicchetti, 1994), or at least "moderate" (Koo & Li, 2016) for exploratory studies.

However, it is worth noting that the relatively low ICC may partly reflect a deliberate design choice. The rubric was intentionally open-ended, relying on each rater's professional judgment as a proficient English speaker to evaluate what constitutes acceptable task performance. This was based on the premise that standards of communicative success often vary across interlocutors in real-world interactions with second language users. Nevertheless, the ambiguity in some rubric items likely contributed to inconsistent interpretations, and future iterations will aim to improve clarity and specificity (e.g., by incorporating behavioral descriptors or anchored examples) to enhance inter-rater agreement while still preserving space for authentic judgment.

The average subtask scores among the raters were used to quantify the extent of transfer using the following formula:

$$\text{Transfer percentage} = \frac{\text{RWT score}}{\text{PT6 score}} \times 100$$

A transfer percentage of 100% indicates complete transfer of task-related language skills to the real-world context.

To supplement these subjective evaluations, conversation analysis further examined the degree of skill transfer by comparing the participants' language use in PT6 and the RWT. As both activities involve bar-ordering tasks with similar rules but differing contexts (virtual vs. the real world), the analysis identifies similarities and changes in the participants' utterances. Evidence of similar language use suggests successful transfer, while negative shifts (such as increased errors or breakdowns in communication)

indicate challenges in applying learned skills in real-world situations. Given the open-ended nature of the RWT, the analysis considers all utterances produced during the tasks.

Conversation analysis used across the pretest, PT6, and the RWT followed a systematic, multi-step procedure. First, all learner–interlocutor interactions were transcribed in full. Second, exchanges involving the target discourse (e.g., making orders, responding to target discourse questions, asking for recommendations) were identified and extracted. Third, the participants' utterances were compared across the three phases to trace developments in the use of task-relevant language.

To provide concrete evidence of changes in learners' language use, a set of consistent coding categories was applied. These included (1) accuracy of target expressions; (2) fluency indicators such as response time, length of pauses, and use of fillers or hesitation markers; and (3) reliance on the first language. These features were selected to reflect both linguistic gains and task completion skills, in alignment with the goals of TBLT. Excerpts presented in the results section have been selected to exemplify observed patterns and developmental trends across participants.

To address the second research question (How do learners perceive their English learning experiences in VR-TBLT?), an exit questionnaire was administered in Japanese. The questionnaire includes items designed to capture the participants' subjective reflections on their VR-TBLT learning experience and perceived transfer to the RWT (see Appendix G). Responses were translated into English and analyzed using In Vivo coding, which involves extracting words or phrases directly from participants' own language (Saldaña, 2016). This method minimizes researcher bias in thematic interpretation. A total of 43 distinct codes were identified and organized into four overarching themes with corresponding subcategories.

Results

Language Gains Through VR-TBLT

Score Comparison

Participant A

The comprehension test included seven question items that asked participants how well they understood the following expressions:

- (1) What can I get for you?
- (2) How's the night going?
- (3) Can I see your ID?
- (4) Do you want to open a tab?
- (5) I could recommend a Manhattan for you.
- (6) How's everything tasting?
- (7) Your total is going to be 50 dollars.

The expression test included eight question items that asked the participants how well they could translate and express the given Japanese phrases in English. The correctly translated expressions are:

- (1) Can I get a glass of beer?
- (2) Do you have any recommendations?
- (3) Do you have any sweet cocktails?

- (4) How strong is that drink?
- (5) What is in the drink?
- (6) How much is that?
- (7) Can I get the bill?
- (8) Keep the change.

Figures 7 and 8 below show the changes in Participant A's self-reported scores regarding comprehension and expression skills in the pretest and the post-test. The numbers on the left side represent Likert-scale options: 1 means "no understanding" and 4 means "full understanding" in the comprehension tests, whereas 1 means "no idea" and 4 means "confident" in the expression tests. The numbers below the bars represent the question numbers.

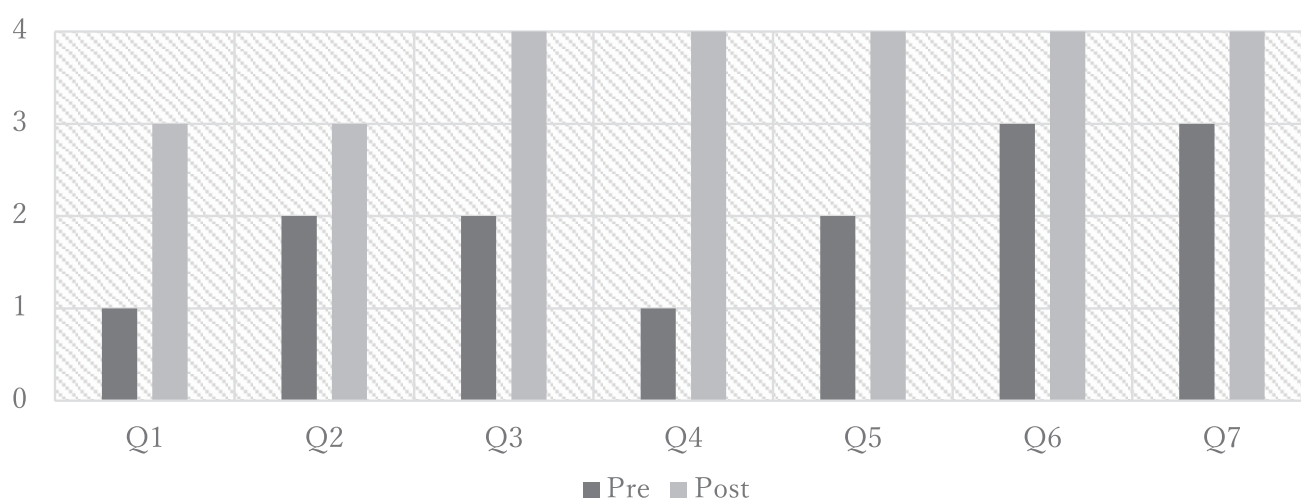


Figure 7 Participant A's Self-Reported Comprehension Skills.

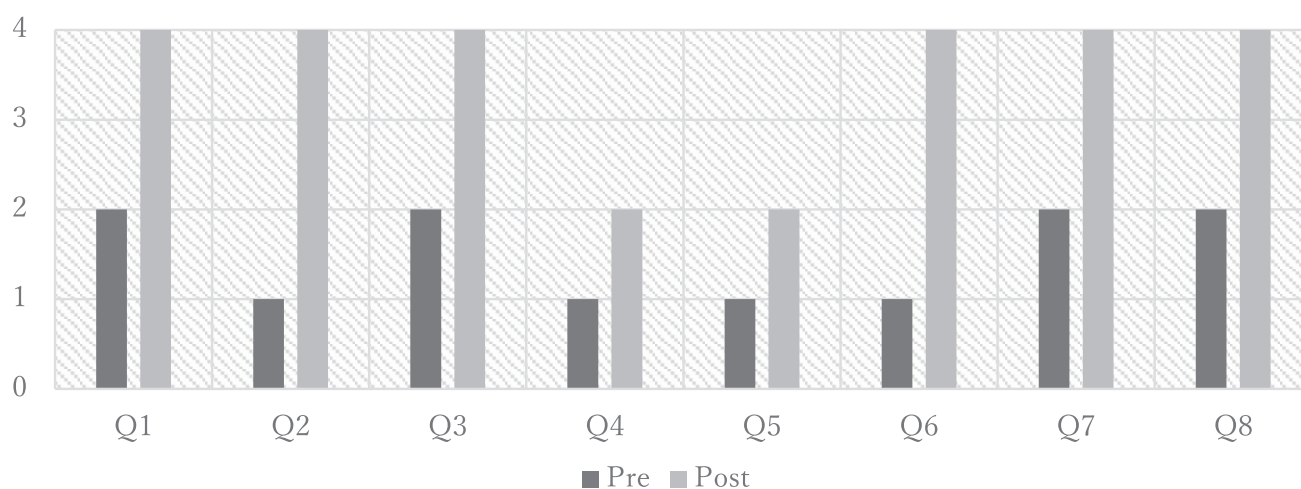


Figure 8 Participant A's Self-Reported Expression Skills.

For the comprehension part, Participant A earned 14 points out of 28 in the pretest, and his score in the post-test was 26 in total. As the chart shows, his comprehension skills showed improvement in every expression. He also reported that his expression skills improved in all the phrases, with his score improving from 12 to 28 out of 32. Overall, Participant A clearly perceives his skills in both understanding and expressing certain expressions as having improved.

Participant B

Participant B's scores appear in Figures 9 and 10 below.

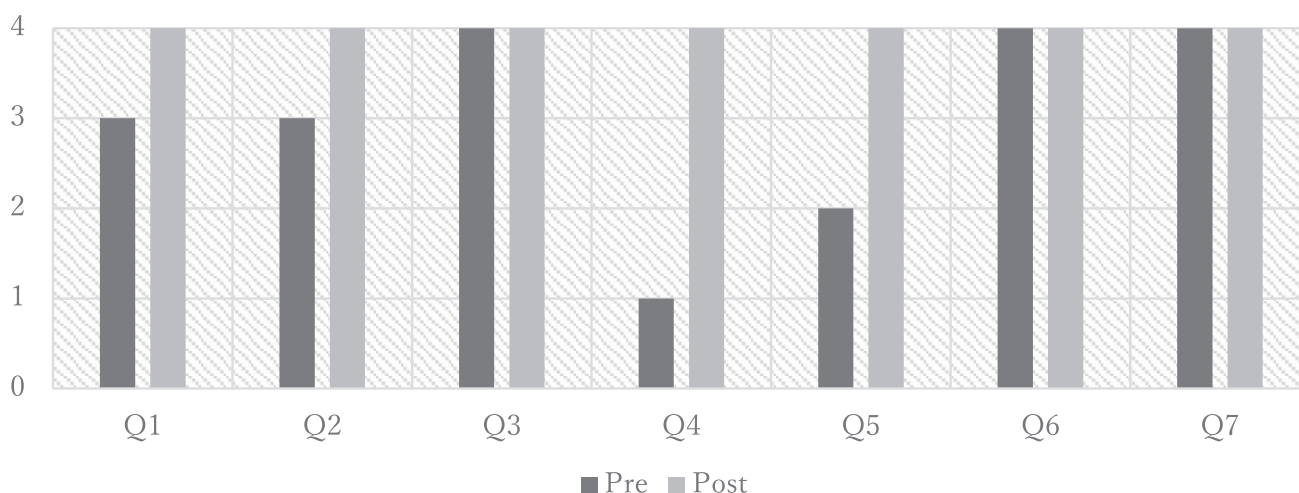


Figure 9 Participant B's Self-Reported Comprehension Skills.

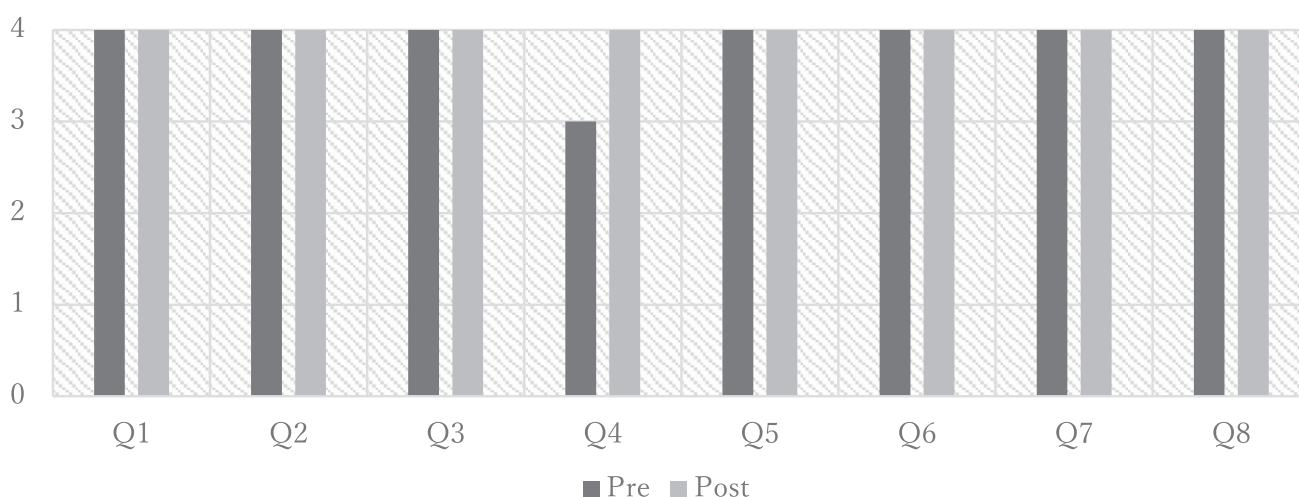


Figure 10 Participant B's Self-Reported Expression Skills.

Participant B, on the other hand, already scored 21 points in the comprehension pretest, which allowed only a 5-point improvement in the post-test, and he reaching the full score of 26 (Table 3). Likewise, his expression skills were found to be very high from the beginning, which did not allow much room for improvement, as he started with 31 and ended with a full score of 32 (Table 4). However, like Participant A, Participant B's overall skills in understanding and expressing certain expressions either improved or were maintained, and they never declined.

Conversation Analysis

This section illustrates how each participant's language use evolved during the intervention, on the basis of conversation analysis of the transcripts. Key improvements are highlighted with excerpts. The abbreviations used are BT (bartender), PA (Participant A), and PB (Participant B). Brackets with a number indicate a silence or pause, showing its duration. For instance, [3.0] represents a 3.0-second

pause. Shorter pauses are represented by “...” as they do not disrupt the conversational flow. In all Excerpts, Japanese expressions are shown in *italics*.

Participant A

One of the most difficult expressions for Japanese learners is “Do you want to start a tab?”, which is rare in Japan and not easily inferred from its individual words. Participant A initially struggled with it, responding with silence. By PT6, however, he responded promptly, as shown below:

Excerpt 1: Response in PT6

BT: So, do you wanna open a tab?

PA: Ah! Yes yes yes yes!

This improvement aligns with Participant A’s self-reported comprehension in the expression tests (Table 1). His ordering skills also improved. Initially relying on simple noun + “please” constructions such as “one beer, please,” by PT6, he used more complete sentences like “Can I have...?” (Excerpt 2) and “Do you have some appetizer?” (Excerpt 3), using the target discourse integrated in the PTs.

Excerpt 2: Ordering drinks in PT6

(He orders the first two drinks)

PA: Uh... [3.5] Can I have beer and a whiskey shot, please?

(When he orders a final drink)

PA: I’ll have... [3.9] Can I have... *kakuteru beeru tanonda* red wine, please?

Excerpt 3: Asking about food in PT6

PA: Ok, *ja*... do you have some appetizer?

Besides target discourse necessary for ordering drinks and food, Participant A acquired the phrase “Keep the change” by PT6, which he had not been able to express during the pretest. He initially hesitated and used first-language interjections. In contrast, Excerpt 4 shows a smooth and confident exchange. This result aligns with his self-evaluation of the corresponding part in the expression test as well (Table 2).

Excerpt 4: Payment in PT6

BT: Its 49 dollars and 50 cents.

PA: Ok... (PA hands BT bills) 55 dollars. Keep the change.

Participant B

While Participant B demonstrated less room for improvement, as suggested by the self-evaluation, conversation analysis revealed subtle gains. His ability to ask for recommendations improved, as seen in Excerpts 5 and 6.

Excerpt 5: Asking for recommendations in the pretest

(When he asks for a drink recommendation)

PB: I like wine, so uh... do you have your any recommendations?

(A few moments later when he asks again)

PB: So, do you have any... do you have any recommends for him, for me?

(When he asks for the third time)

PB: So, uh... besides Manhattan, do you have any recommends?

- Excerpt 6: Asking for recommendations in PT6
 (When he asks for a first drink recommendation)
 PB: Do you have any recommendations?
 (When he orders a third drink)
 PB: I like wine, so, do you have any recommendations?

In the pretest (Excerpt 5), Participant B used unnecessary or incorrect words such as “your” and “recommends.” By PT6 (Excerpt 6), he fluently used “Do you have any recommendations?” as a complete formulaic expression. He also demonstrated improvement in placing orders, using another target expression in PT6:

- Excerpt 7: Placing an order in PT6
 PB: I want something um... sweet one, sweet cocktail. Do you have any idea?
 BT: How about a Margarita?
 PB: Ah, I'll do this, I'll do it.
 BT: All right, coming right up.

The phrase “I'll do...” is very simple, but not many Japanese students know that they can use it to make orders (e.g., “I'll do a Manhattan.”). The phrase “I'll do...” was used by a native speaker in the modeling videos during PTs, and Participant B seems to have learned the phrase through those videos and tried it out in role-plays.

Transfer of Skills from VR-TBLT to RWT

Rubric-Based Rating

The rubric has six subtasks to evaluate per test using a 10-point Likert scale. Two participants were assessed in PT6 and RWT, making 24 subtasks to evaluate in total. The averaged scores among the raters were used to produce transfer rates for each subtask with the formula introduced earlier. The six subtasks consist of: (1) showing IDs, (2) first orders, (3) second orders, (4) tipping, (5) asking for recommendations, and (6) improvising an unexpected order, as shown in Appendix H.

The resulting transfer rates for the six subtasks regarding Participant A's performance are (1) 118%, (2) 105%, (3) 114%, (4) 87%, (5) 105%, and (6) 96%, making the average transfer rate 104%. The transfer rates for Participant B are (1) 104%, (2) 112%, (3) 112%, (4) 145%, (5) 112%, and (6) 112%, with an average rate of 116%. Transfer rates exceeding 100% indicate that both participants transferred their performance skills acquired in the PTs to RWT very well and they even improved their performance.

Conversation Analysis

This section analyzes each participant's transfer of the target discourse learned during PTs, using conversation analysis of the transcripts. Similar to the conversation analysis exploring language gains, prominent instances of transfer are highlighted with concrete examples.

Participant A

The first set of transfer examples appears in Excerpt 8, in which Participant A orders drinks in the RWT.

- Excerpt 8: Ordering drinks in the RWT
 PA: Uh... [5.6] (looking at the drink menu) Can I have, uh, can I have a non-alcohol beer?

- BT: Ah, okay.
PA: He is... uh... [13.6] Ah, do you have something to beer?
BT: Oh, yeah, some beer? Okay, yeah.
(A few moments later when he orders another drink for himself)
PA: And, uh... do you have something to non-alcohol cocktail?
BT: Ah, yes, I can.
(When his friend requests a sweet cocktail in Japanese)
PA: Uh, do you have something uh... [2.1] sweet cocktails?
BT: Yes, yeah.

The first notable transfer is the question structure “Can I have...?”, which Participant A acquired during treatments, as shown previously in Excerpt 2. Additionally, Excerpt 8 contains three instances of “Do you have...?”, another target discourse item he displayed in PT6. Despite some grammatical errors, Participant A seems to have internalized “Do you have...?” as a chunk and is attempting to adapt it to create new sentences conveying his intentions. He also successfully transfers target discourse items from Excerpts 3 and 4, as illustrated in Excerpts 9 and 10.

- Excerpt 9: Asking about food in the RWT
PA: Uh... do you have some appetizers?
BT: Oh, yes, we do, yeah.

- Excerpt 10: Making payment in the RWT
BT: So, your total will be 42 dollars.
PA: Okay. (Counting the bills) Here you are.
BT: Thank you.
PA: 47 dollars.
BT: Ah, 47.
PA: Keep the change.
BT: Oh, thank you so much.

Participant B

One of Participant B’s successful transfers is the phrase used to ask for recommendations. As noted in earlier conversation analyses, his broken expression in the pretest improved in PT6 (Excerpt 6). Excerpt 11 demonstrates his transfer of this expression to the RWT. Interestingly, as in the second half of Excerpt 11, he adapts the phrase to discuss places to visit in the U.S.

- Excerpt 11: Asking for recommendations in the RWT
(When he hears that his friend wants a local beer)
PB: He says, uh, he, uh, he likes to, uh, drink, uh, something beer, a local beer.
BT: Okay.
PB: Do you have any recommendations?
(When he has a random talk with the bartender)
PB: I want to visit America in the future, so, do you have any recommendation to visit?
BT: Ah, let’s see. I really like Los Angeles...

Excerpt 12 highlights Participant B’s natural use of the phrase “I’ll do...”, which he successfully transfers multiple times to the RWT.

Excerpt 12: Placing an order in the RWT

PB: So, what is the shot, like whiskey, gin?

BT: So, it's very common, especially in Montana, to drink whiskey. So, a lot of people drink bourbon.

PB: Oh, bourbon. I'll do it.

(Later when he orders another drink for himself)

PB: (Pointing at the shot of bourbon) This so strong, so, I'll do non-alcohol.

Participants' Perceptions of the VR-TBLT

Participants' responses to the questionnaire were analyzed through In Vivo coding, and the resulting codes were categorized into four themes with several subcategories. Each theme is presented in tables (Tables 1–4), and key elements in the participants' answers are explained with examples.

Theme 1: Advantages of Learning in VR

Table 1 *Advantages of Learning in VR*

Subcategories	In Vivo codes
Effectiveness of learning activities	"Activities were very useful." "Everything was useful." "Activities were very useful." "Adds positive pressure."
Realistic and immersive experiences in VR	"As if it hadn't been my first time." "Able to experience studying abroad in advance." "In a setting very close to reality." "Learn by immersing yourself in VR." "You feel immersed." "As if you were in a real bar." "You can learn through realistic experiences." "It's... realistic."
Recommendations for incorporating VR into language learning	"Suggest using VR." "You can learn a lot." "You get more learning takeaways from VR learning."

Both participants highlighted the practicality of VR's task-based activities. For example, Participant A stated, "Everything was useful, like learning a variety of vocabulary, the tipping culture, and many other things." This underscores the comprehensive utility of the VR-TBLT tasks.

Participants frequently emphasized the realism of VR, noting its ability to simulate real-life experiences, such as being in a bar. This immersive quality enabled them to practice language in highly authentic environments, preparing them for real-world scenarios. Participant A commented, "I felt in real-world tasks as if it hadn't been my first time," while Participant B remarked, "I was able to experience studying abroad in advance," implying that the learning experience was realistic enough to serve as a rehearsal for real-life interactions.

At the end of the questionnaire, both participants compared their VR learning experience with traditional methods and strongly recommended VR as an effective language learning tool. For instance, Participant A said, "I will definitely suggest using VR. You can learn a lot," while Participant B stated, "I also feel that you get more learning takeaways from VR learning than from traditional lessons."

Theme 2: Learning Outcomes from VR-Based Learning

Table 2 *Learning Outcomes from VR-Based Learning*

Subcategories	In Vivo codes
Perceived learning achievements	"My knowledge expanded." "Learned the entire process of ordering in a bar." "Able to learn bar-related vocabulary."
Awareness of skill transfer to real-life scenarios	"Able to make myself understood." "Able to complete the task." "Able to use what I learned."

The participants achieved significant learning outcomes, including linguistic and non-linguistic outcomes. Participant B, for example, emphasized this by saying, "I learned the entire process of ordering in a bar," and "I was able to learn bar-related vocabulary in a setting very close to reality."

Regarding transfer of learning, both participants expressed self-awareness of their ability to apply language and task-completion skills in real-life situations. Participant A reflected, "I was able to complete the task," while Participant B stated succinctly, "I was able to use what I learned."

Theme 3: Positive Impacts of VR-Based Learning on Learners' Emotions

Table 3 *Positive Impacts of VR-Based Learning on Learners' Emotions*

Subcategories	In Vivo codes
Enjoyment derived from the VR experience	"Fun to learn them." "It was enjoyable." "VR was enjoyable." "As if we were playing video games." "VR learning is far more enjoyable than classroom learning." "It's also enjoyable." "You don't get bored." "It's enjoyable."
Increased motivation through VR-based learning	"Felt like trying them out." "Felt motivated." "Would like to try other versions of VR."
Reduced anxiety and boosted confidence	"I would have been more nervous without learning in VR." "No problem with English conversation."

Enjoyment emerged as a recurring theme across both participants. They consistently described VR as engaging and fun, with an entertaining experience resembling video games. Participant A noted, "VR was enjoyable above all. It's nice that we can learn as if we were playing video games." Similarly, Participant B highlighted the advantages of VR over traditional classroom methods, stating, "VR learning is far more enjoyable than classroom learning," and adding, "You don't get bored. It's enjoyable and realistic."

The immersive and interactive nature of VR significantly enhanced the participants' motivation to learn and explore further. Participant A, in particular, expressed high levels of motivation, stating, "As my knowledge expanded, I felt like trying them out," and "I am not good at studying and personally do not like it, but I felt motivated. I would like to try other versions of VR learning as well."

Additionally, Participant A highlighted VR's capacity to reduce anxiety by providing a realistic environment for practicing language skills. He reflected, "I was nervous, but I think I would have been more nervous without learning in VR." Participant B demonstrated strengthened confidence in his language abilities, stating, "I think I will have no problem with English conversation in bars."

Theme 4: Challenges of Learning in VR

Table 4 *Challenges of Learning in VR*

Subcategories	In Vivo codes
None	"Not able to understand when I started." "Tough to learn a lot of things." "Hard to remember words." "Forgot some things." "I was very nervous." "I was nervous." "Vocabulary... was difficult... in listening exercises." "Hard to operate the VR equipment."

However, both participants acknowledged challenges in learning through VR. These included initial struggles with understanding and retaining vocabulary, nervousness during tasks, and technical difficulties with the VR equipment. Participant A commented, "It was hard to remember words. Unlike rote memorization, I had to understand the meanings behind them," suggesting that context-rich VR learning requires a deeper understanding of language for practical use. Participant B, meanwhile, pointed to technical challenges, saying, "It was a little hard to operate the VR equipment."

Discussion

This section addresses the two research questions that guided the study:

1. To what extent can a custom VR-TBLT program contribute to the acquisition of task-related language skills?
2. How do learners perceive their English learning experiences in a VR-TBLT context?

RQ1: Language Skill Acquisition and Transfer

Findings from both of the self-evaluation tests and conversation analysis suggest that the VR-TBLT program supported participants' acquisition and use of task-related language. Despite the small sample, both participants demonstrated improved use of the target discourse across tasks. These outcomes align with prior studies highlighting learners' favorable perceptions of VR for second language acquisition (e.g., Alizadeh & Cowie, 2022).

Crucially, participants developed these skills without explicit grammar instruction, relying instead on modeling videos and task handouts—hallmarks of a task-based approach. The repeated, communicative use of formulaic expressions appeared especially effective for transfer, consistent with Gatbonton and Segalowitz's (1988) criteria for successful formulaic language use.

Transfer was observed from the VR tasks to the real-world transfer test (RWT). Some transfer rates exceeded 100%, which may reflect the rehearsal effect noted by Ellis (2009), particularly given the high similarity between the VR and real-world settings. Although some may question whether this constitutes "meaningful" transfer, research on transfer theory (e.g., Detterman, 1993) highlights the value

of near transfer, especially in language learning. However, the observation that one participant flexibly adapted newly learned expressions to new pragmatic contexts also supports the view of transfer as a dynamic and adaptive process (Larsen-Freeman, 2013), not mere repetition in the same contexts.

RQ2: Learners' Perceptions

The participants' perceptions, analyzed through In Vivo coding of the questionnaire responses, were overwhelmingly positive. The most salient theme was "enjoyment of the VR experience," reflecting findings from other studies (Yudintseva, 2023; Zheng et al., 2022). While this may partly reflect the novelty effect (Smith & McCurrach, 2021), it nonetheless indicates the motivational potential of VR in language learning.

The second prominent theme was the "realism and immersion" provided by the VR environment. Participants reported that they "felt like they were actually at a bar," suggesting a strong sense of presence, one of the distinctive affordances of VR. Prior studies have linked increased presence in VR with enhanced engagement and learning outcomes (Essoe et al., 2022; Sun et al., 2020), and these findings are echoed here.

At the same time, participants also reported several "challenges of learning in VR," consistent with previous research (Cowie, 2022; Ding, 2024). These included technical difficulties, such as manipulating virtual objects and reading in-VR handouts, as well as physical discomfort like dizziness. One participant overcame such issues by shifting to real-body movement, underscoring the importance of design flexibility. These challenges reinforce the need for close collaboration between educators and developers when implementing VR in language education.

Limitations and Future Directions

As an exploratory case study with only two participants—both male and of differing proficiency levels—the findings are not generalizable. Participant B, who was more proficient initially, showed smaller gains, suggesting that baseline proficiency may mediate the effects of task complexity. As suggested by Khezrlou (2025), future research should investigate this interaction with larger, more diverse samples.

Moreover, the improvements observed in this study cannot be solely attributed to the use of VR. The positive outcomes may be due to the task-based framework itself, which has demonstrated effectiveness even in non-immersive contexts (e.g., Benson, 2016; Nolen, 2020). To isolate the unique contributions of VR, future studies should use comparative designs contrasting VR-TBLT with non-VR TBLT, ideally under otherwise equivalent pedagogical conditions.

Finally, the relatively low inter-rater reliability in task performance assessment signals the need for clearer, more behaviorally anchored rubrics to enhance consistency and interpretability in future work.

Conclusion and Implications for Teachers

This study responded to recent calls for pedagogically grounded VR content by offering a custom-designed VR-TBLT program that aligns with the core principles of TBLT. Although limited in scale, the study contributes a model of how educators can design immersive, goal-oriented language tasks in VR environments.

The findings from this small case study suggest that, in this particular context, the two participants were able to acquire and transfer task-relevant language without explicit instruction. This appeared

to occur through repeated, communicative use of the target expressions in the VR environment. Both participants also reported strong motivation and engagement, which seemed to be supported by the realism and interactivity of the VR setting. While these results cannot be generalized, they offer preliminary support for the potential of VR to provide a safe space for rehearsal and experimentation, particularly for learners who may benefit from immersive, low-pressure practice opportunities.

For teachers interested in incorporating VR into task-based courses, the following practical insights emerge:

- Explore the process of creating custom-designed VR-TBLT materials—even without technical expertise—by using free tools and resources.
- Replicate real-world settings where learners are likely to use the target language.
- Embed repeated use of formulaic expressions tied to meaningful subtasks.
- Prepare for technical challenges by offering alternatives (e.g., physical navigation, printed hand-outs outside VR, training sessions).

The immersive and interactive nature of VR, combined with the pedagogical strength of TBLT, may create fertile ground for robust learning of target discourse, including the potential for transfer to real-world contexts. Future research should explore whether and how VR-TBLT can foster such broader, adaptive language use.

Conflicts of Interest

The author declares no conflicts of interest.

Use of Generative AI

I would like to acknowledge the use of ChatGPT-4.0, an AI language model developed by OpenAI, for its assistance in revising the English in this paper. Any remaining errors are solely the author's responsibility.

References

- Alizadeh, M., & Cowie, N. (2022). Language learning and virtual reality: A scoping review. Reconnecting relationships through technology. *ASCILITE Publications*, 13, e22258. <https://doi.org/10.14742/apubs.2022.258>
- Benson, S. D. (2016). Task-based language teaching: An empirical study of task transfer. *Language Teaching Research*, 20(3), 341–365. <https://doi.org/10.1177/1362168815569829>
- Christoforou, M., Thrasher, T., & Ijiri, A. (2024). Technology-mediated task-based language teaching in a virtual reality learning environment. In Y. Choubasaz, P. Díez-Arcón, A. Gimeno-Sanz, J. Hriňák, X. Liashuk, S. Pokrivčáková & H. Vančová (Eds.), *CALL for humanity – EuroCALL 2024 short papers*. <https://doi.org/10.4995/EuroCALL2024.2024.19028>
- Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment*, 6(4), 284. <https://doi.org/10.1037/1040-3590.6.4.284>
- Ciekanski, M., Kalyaniwala, C., Molle, N., & Provas-Bréauté, V. (2020). Real and perceived affordances of immersive virtual environments in a language teacher-training context: Effects on the design of learning tasks. *Revista Docência e Cibercultura*, 4(3), 83–111. <https://doi.org/10.12957/redoc.2020.56752>

- Cowie, N. (2022). The affordances and challenges of virtual reality for language teaching. *International Journal of TESOL Studies*, 4(3), 50–65. <https://doi.org/10.46451/ijts.2022.03.05>
- Detterman, D. K. (1993). The case for the prosecution: Transfer as an epiphenomenon. In D. K. Detterman & R. J. Sternberg (Eds.), *Transfer on trial: Intelligence, cognition, and instruction* (pp. 1–24). Ablex Publishing.
- Dhimolea, T., Kaplan-Rakowski, R., & Lin, L. (2022). A systematic review of research on high-immersion virtual reality for language learning. *TechTrends*, 66, 810–824. <https://doi.org/10.1007/s11528-022-00717-w>
- Ding, M. (2024). The impact of high-immersion virtual reality on EFL learners' foreign language speaking anxiety: A mixed-method approach. *ReCALL*, 36(3), 287–305. <https://doi.org/10.1017/S0958344024000156>
- Ellis, R. (2003). *Task-based language learning and teaching*. Oxford University Press.
- Ellis, R. (2009). The differential effects of three types of task planning on the fluency, complexity, and accuracy in L2 oral production. *Applied Linguistics*, 30(4), 474–509. <https://doi.org/10.1093/applin/amp042>
- Essoe, J., Reggente, N., Ohno, A., Beak, Y., Dell'Italia, J., & Rissman, J. (2022). Enhancing learning and retention with distinctive virtual reality environments and mental context reinstatement. *NPJ Science of Learning*, 7(31), 1–14. <https://doi.org/10.1038/s41539-022-00147-6>
- Frazier, E., Bonner, E., & Lege, R. (2018). A brief investigation into the potential for virtual reality: A tool for 2nd language learning distance education in Japan. *The Language and Media Learning Research Center Annual Report, 2018*, 213–218. <https://kuis.repo.nii.ac.jp/records/1789>
- Gatbonton, E., & Segalowitz, N. (1988). Creative automatization: Principles for promoting fluency within a communicative framework. *TESOL Quarterly*, 22(3), 473–492. <https://doi.org/10.2307/3587290>
- Hoang, D., Johnson, N., & McAlinden, M. (2023). Exploring the potential of VR in enhancing authentic learning for EFL tertiary students in Vietnam. *Teaching English with Technology*, 23(1), 1–22.
- Iizuka, T. (2019). Task-based needs analysis: Identifying communicative needs for study abroad students in Japan. *System*, 80, 134–142. <https://doi.org/10.1016/j.system.2018.11.005>
- Jackson, D. O. (2022). *Task-based language teaching*. Cambridge University Press. <https://doi.org/10.1017/9781009067973>
- Kaplan-Rakowski, R., & Gruber, A. (2019). *Low-immersion versus high-immersion virtual reality: Definitions, classification, and examples with a foreign language focus* [Conference session]. Proceedings of the Innovation in Language Learning International Conference 2019, Pixel, 13, SLA4232. <https://conference.pixel-online.net/files/ict4ll/ed0012/FP/6143-SLA4232-FP-ICT4LL12.pdf>
- Kawasumi, S., & Ishii, Y. (2023). Comparative study of English learning using virtual reality and a smartphone application. *Bulletin of the Faculty of Education, Chiba University*, 71, 99–105. <https://opac.ll.chiba-u.jp/da/curator/900121724/>
- Khezrlou, S. (2025). Task-based language teaching. In Y. Wang & G. Stockwell (Eds.), *The Cambridge handbook of technology in language teaching and learning* (pp. 377–392). Cambridge University Press. <https://doi.org/10.1017/9781009294850.028>
- Kim, Y., & Namkung, Y. (2024). Methodological characteristics in technology-mediated task-based language teaching research: Current practices and future directions. *Annual Review of Applied Linguistics*. Published online ahead of print. <https://doi.org/10.1017/S0267190524000096>
- Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine*, 15(2), 155–163. <https://doi.org/10.1016/j.jcm.2016.02.012>

- Kostopoulou, S. (2009). Democratising and enhancing the quality of institutionalised language assessment through the European Language Portfolio. In L. Taylor & C. J. Weir (Eds.), *Language testing matters: Investigating the wider social and educational impact of assessment* (pp. 263–278). Cambridge University Press.
- Larsen-Freeman, D. (2013). Transfer of learning transformed. *Language Learning*, 63(1), 107–129. <https://doi.org/10.1111/j.1467-9922.2012.00740.x>
- Nolen, C. (2020). *Examining task transferability in task-based language teaching: A multi-case study* [Doctoral dissertation, Georgia State University]. ScholarWorks @ Georgia State University. https://scholarworks.gsu.edu/alesl_diss/53
- Nunan, D. (2010). A task-based approach to materials development. *Advances in Language and Literary Studies*, 1(2), 135–160. <http://dx.doi.org/10.7575/aiac.all.v.1n.2p.135>
- Oda, M. (2024). University teachers' attitudes towards the introduction of VR in foreign language learning: Based on technological pedagogical content knowledge (TPACK). *Research Report of JET Conferences*, 2024(2), 111–116. https://doi.org/10.15077/jsetstudy.2024.2_111
- Ondarra, J., K., Melchor-Couto, S., & Canto, S. (2025). Virtual reality. In Y. Wang & G. Stockwell (Eds.), *The Cambridge handbook of technology in language teaching and learning* (pp. 377–392). Cambridge University Press. <https://doi.org/10.1017/9781009294850.028>
- Parmaxi, A. (2020). 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>
- Saldaña, J. (2016). *The coding manual for qualitative researchers* (3rd ed.). Sage.
- Shintani, N. (2018). Researching TBLT for young, beginner learners in Japan. In M. Bygate (Ed.), *TBLT as a researched pedagogy* (pp. 199–212). John Benjamins. <https://doi.org/10.1075/tblt.12.08shi?locatt=mode:legacy>
- Smith, M., & McCurrach, D. (2021). The usage of virtual reality in task-based language teaching. *Proceedings of the 28th Korea TESOL International Conference: Re-envisioning ELT Altogether, All Together, KOTESOL*, 28(1), 153–165. <https://koreatesol.org/content/kotesol-proceedings>
- Sun, C., Yao, Y., Wang, R., & Ye, X. (2020). A study of the influence of scene reality of VR environment on English learners' learning engagement and learning effectiveness [Conference session]. *2020 IEEE 2nd International Conference on Computer Science and Educational Informatization (CSEI)* (pp. 181–185). <https://doi.org/10.1109/CSEI50228.2020.9142520>
- Taguchi, N., & Zhao, S. (2025). Can virtual reality leverage technology-mediated task-based language teaching? A research synthesis. *Digital Studies in Language and Literature*, 2(1), 25–53. <https://doi.org/10.1515/dsll-2024-0028>
- Wiboolyasarini, W., Jinowat, N., Wiboolyasarini, K., Kamonsawad, R., Tiranant, P., & Boonyakitanont, P. (2023). Enhancing L2 speaking proficiency through collaborative tasks in RILCA world: the case of East Asian learners. *Asian-Pacific Journal of Second and Foreign Language Education*, 8, 37. <https://doi.org/10.1186/s40862-023-00209-1>
- Yamanaka, T., Miyazaki, Y., Yamazaki, A., Murakami, K., Kimura, S., Yamashita, M., & Kondo, Y. (2023). Testing virtual reality for eliminating Japanese University students' English-speaking anxiety: Cases of international conference and restaurant. *International Journal of Learning and Teaching*, 9(1), 49–55. <https://doi.org/10.18178/IJLT.9.1.49-55>
- Yan, W., & Lowell, V. L. (2024). Design and evaluation of task-based role-play speaking activities in a VR environment for authentic learning: A design-based research approach. *The Journal of Applied Instructional Design*, 13(4), 309–327. <https://doi.org/10.59668/2033.18647>
- Yano, K. (2022). Potential of VR and Metaverse in foreign language education. *Computer & Education*, 53, 37–42. <https://doi.org/10.14949/konpyutariyoukyouiku.53.37>

- Yudintseva, A. (2023). Virtual reality affordances for oral communication in English as a second language classroom: A literature review. *Computers & Education: X Reality*, 2, 100018. <https://doi.org/10.1016/j.cexr.2023.100018>
- Zheng, C., Yu, M., Guo, Z., Liu, H., Gao, M., & Chai, C. (2022). Review of the application of virtual reality in language education from 2010 to 2020. *Journal of China Computer-Assisted Language Learning*, 2(2), 299–335. <https://doi.org/10.1515/jccall-2022-0014>

Appendix A

Handout for PT1

PT1: Find out an ordering procedure at a bar in the US.

1. Watch a video in which a man orders a drink in a bar in the US. You can watch it as many times as you want.
2. Focus on how American people make orders in a bar, looking at the list below.
3. Fill in the blanks in the list.
 - The bartender asks the customer to show his ().
 - The bartender asks the customer whether to open a () or not.
 - The customer asks for recommendations and tells the bartender his preference.
 - The bartender makes and serves a drink for him.
 - When the customer pays for his drink, he gives the bartender some extra money as a ().
 - After finishing his drink, the customer leaves the bar.

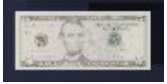
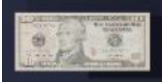
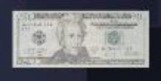
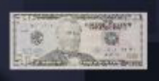
Appendix B

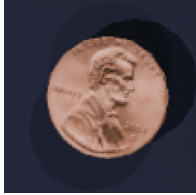
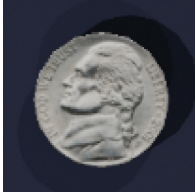
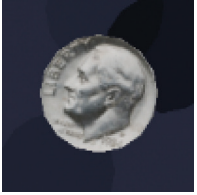
Handout for PT2

PT2: Find out about US bills and coins.

1. Watch a video in which a man talks about US bills and coins. You can watch it as many times as you want.
2. Learn how to identify each bill and coin, focusing on who is printed on them.
3. Complete the tables below by choosing appropriate names from the glossary box.
4. After completing the table, practice paying certain amounts of money using US bills and coins in VR.

Who's on American money?

BILLS						
	1 dollar	5 dollars	10 dollars	20 dollars	50 dollars	100 dollars
WHO						

COINS				
	1 cent	5 cents	10 cents	25 cents
WHO				
NICKNAMES	Penny	Nickel	Dime	Quarter

Glossary:

1. Thomas Jefferson / 2. Alexander Hamilton / 3. Ulysses S. Grant / 4. Andrew Jackson / 5. Abraham Lincoln / 6. George Washington / 7. Franklin D. Roosevelt / 8. Benjamin Franklin

Appendix C

Handout for PT3

PT3: Find out about alcoholic drinks and how to describe them.

1. Make a pair.
2. One learner gets an A handout and the other gets a B handout.
3. Interact with each other to figure out what goes into the blanks in the handouts.

A handout:

How much alcohol does each drink contain?

beer	wine	whiskey	tequila	gin	rum	vodka
5%	()%	40%+	()%+	40%+	()%+	40%+

Useful phrases and expressions!

(ビール) のアルコール度数はどれくらいですか？

How much alcohol is in (beer)?

How much alcohol does (beer) have?

How strong is (beer)?

What's the percentage of (beer)?

(5) %くらい/以上です。

It has about/more than (5)% alcohol content.

It contains about/more than (5)% alcohol.

It is about (5)% alcohol.

Appendix D

Handout for PT4

PT2: Find out about alcoholic drinks and how to describe them.

1. Make a pair.
2. One learner gets an A handout and the other gets a B handout.
3. Interact with each other to figure out what goes into the blanks in the handouts.

A Handout:

Ask about the cocktails below and find out what goes into the blanks.

	BASE	MIX	GARNISH	TASTE
Manhattan 	bourbon or whiskey	> sweet vermouth > Angostura bitters	cherry	bitter and slightly sweet
Negroni 				bitter, sour, and slightly sweet
Old-fashioned 	bourbon or whiskey	> sugar cube > Angostura bitters	cherry or orange	sweet, bitter, and spicy
Margarita 		> triple sec > lime juice > salt		
Whiskey sour 	bourbon	> lemon juice > syrup > egg white	cherry and orange	sweet and sour
The martini 	gin or vodka	> dry vermouth	olive or lemon	bitter, sweet, dry, and salty

Useful phrases and expressions!

（マンハッタン）とは何ですか。 / （マンハッタン）はどんなカクテルですか。

What is a (Manhattan)? / What kind of cocktail is a (Manhattan)?

（ウイスキー）と（スイートベルモット）と（アンゴスチュラビターズ）で作り、（サクラambo）で飾ります。

It's made with (whiskey), (sweet vermouth), and (Angostura bitters). It's garnished with a (cherry).

We pour (whiskey), add (sweet vermouth) and (Angostura bitters). It's garnished with a (cherry).

どんな味ですか。 / （苦）くて、（少しだけ甘い）味です。

What does it taste like? / It tastes (bitter) and (slightly sweet).

Appendix E

Handout for PT5

PT2: How much should you tip a bartender?

It is said that people usually tip around 20%. However, exactly how much they pay depends on the individual. Some people tip servers \$1 for each simple drink like beer or wine and \$2 for more complex drinks like cocktails. Others pay for drinks with bills and tell the server to keep the change.

1. If you were ordering the following drinks, how much would you pay?

	PRICE	TIP	TOTAL
A whisky shot	\$5.50	\$	\$
A glass of beer	\$7.00	\$	\$
A glass of beer	\$8.50	\$	\$
A gin cocktail	\$10.00	\$	\$
A vodka cocktail	\$11.50	\$	\$
A glass of wine	\$12.00	\$	\$
A plate of appetizers	\$13.50	\$	\$

2. Mini role play

1. Make a pair.
2. One learner acts as a bartender and the other acts as a customer.
 <bartender> Serve your partner drinks and find out how much your partner tips you as they pay for each drink. Fill in the blanks below.
 <customer> Get each drink and tip your partner.
3. Change the roles of bartender and customer.
4. Compare your tips with your partner and discuss how much we should pay.

How much did your partner tip?

	PRICE	TIP	TOTAL
A whiskey shot	\$5.50	\$	\$
A glass of beer	\$7.00	\$	\$
A glass of beer	\$8.50	\$	\$
A gin cocktail	\$10.00	\$	\$
A vodka cocktail	\$11.50	\$	\$
A glass of wine	\$12.00	\$	\$
A plate of appetizers	\$13.50	\$	\$

Useful expressions!

こちらが（ウイスキーのショット）でございます。

Here is your (whiskey shot). / This is your (whiskey shot). Here you are.

お釣りは必要でしょうか。Do you need change?

いいえ、必要ありません。とっておいてください。

No, thank you. / Keep the change.

（ 3ドル ）だけ返してもらえますか。

Please give me (\$3) back. / Can I get (\$3) back?

Appendix F

Task Setting and Steps

Setting:

You are visiting Montana in the United States for sightseeing with your Japanese friend. Tonight, you are going to a bar together. However, your friend cannot speak English at all. Therefore, you want to make orders in English for him instead. You have a \$85 budget: two \$20 bills, three \$10 bills, two \$5 bills, and five \$1 bills.

Steps:

1. Enter the bar and say “hello” to the bartender to start the task.
2. Talk to the bartender and order two drinks (for you and your friend).
3. Talk to the bartender and order an appetizer.
4. Talk to the bartender and order another two drinks (for you and your friend).
5. Say “bye” to the bartender to finish the task.

Instructions (rules):

- Your orders need to include at least one drink recommended by the bartender.
- You cannot order the same drink twice.
- The two orders for your friend need to be based on their preferences.

Appendix G

Questions Asked in the Exit Questionnaire

1. How much do you think your bar-related English has improved?
2. How hard was it to learn bar-related English?
3. What was the most difficult?
4. How useful was virtual reality? Why?
5. How useful were the tasks and activities? Why?
6. What was the most useful?
7. What was your favorite part of the learning process?
8. Were you able to transfer your learning about the ordering process to the real bar? Why?
9. Were you able to transfer your learning of language expressions to the real bar? Why?
10. How different did you feel the learning English in VR was compared with normal English education?
11. Do you think English education in VR is more effective than English education in normal classrooms? Why?
12. Do you think English education in VR is more enjoyable than English education in normal classrooms? Why?

Appendix H

Task Performance Rubric Sheet

Task performance rubric:

By the end of this project, learners will be able to understand how a bar system works in the United States and how to make orders there. To evaluate the learner's task performance, mark an appropriate number that best describes the learner's performance in reference to the following criteria.

Criteria

- 0 – 1 (Poor): Many critical problems and failed to complete the task
 2 – 4 (Unacceptable): A lot of problems and difficulties to complete the task
 5 – 8 (Acceptable): Not perfect, but good enough to complete the task
 9 – 10 (Perfect): Successful performance without any critical problems

SUBTASKS				
1. Identification Success: the learner shows his ID when the bartender requires it.				
10 --- 9 --- 8 --- 7 --- 6 --- 5 --- 4 --- 3 --- 2 --- 1 --- 0				
Perfect	Acceptable	Unacceptable	Poor	
2. Orders 1 (drinks) Success: the learner talks to the bartender and orders four different drinks.				
10 --- 9 --- 8 --- 7 --- 6 --- 5 --- 4 --- 3 --- 2 --- 1 --- 0				
Perfect	Acceptable	Unacceptable	Poor	
3. Orders 2 (food) Success: the learner asks the bartender for a food menu and orders food from the given options.				
10 --- 9 --- 8 --- 7 --- 6 --- 5 --- 4 --- 3 --- 2 --- 1 --- 0				
Perfect	Acceptable	Unacceptable	Poor	
4. Tips Success: the learner's payments include acceptable amounts of tips (more than 10%) to the bartender.				
10 --- 9 --- 8 --- 7 --- 6 --- 5 --- 4 --- 3 --- 2 --- 1 --- 0				
Perfect	Acceptable	Unacceptable	Poor	
5. Recommendation Success: the learner asks the bartender about their recommendation and orders at least one recommended drink.				
10 --- 9 --- 8 --- 7 --- 6 --- 5 --- 4 --- 3 --- 2 --- 1 --- 0				
Perfect	Acceptable	Unacceptable	Poor	
6. Request Success: the learner orders appropriate drinks based on their friend's preferences.				
10 --- 9 --- 8 --- 7 --- 6 --- 5 --- 4 --- 3 --- 2 --- 1 --- 0				
Perfect	Acceptable	Unacceptable	Poor	