THESIS

LEARNING TECHNICAL SPANISH WITH VIRTUAL ENVIRONMENTS

Submitted by

Caspian Siebert

Department of Computer Science

In partial fulfillment of the requirements

For the Degree of Master of Science

Colorado State University

Fort Collins, Colorado

Fall 2024

Master's Committee:

Advisor: Francisco R. Ortega

Alyssia Miller De Rutté
Nikhil Krishnaswamy

ABSTRACT

LEARNING TECHNICAL SPANISH WITH VIRTUAL ENVIRONMENTS

As the world becomes increasingly interconnected through the internet and travel, foreign language learning is essential for accurate communication and a deeper appreciation of diverse cultures. This study explores the effectiveness of a virtual learning environment employing Artificial Intelligence (AI) designed to facilitate Spanish language acquisition among veterinary students in the context of diagnosing a pet. Students' engagement with virtual scenarios that simulate real-life veterinary consultations in Spanish is examined using a qualitative thematic analysis. Participants have conversations with a virtual pet owner, discussing symptoms, diagnosing conditions, and recommending treatments, all in Spanish. Data was collected through recorded interactions with the application and a semi-structured interview. Findings suggest that immersive virtual environments enhance user engagement and interest, and several suggestions were made to improve the application's features. The study highlights the potential for virtual simulations to bridge the gap between language learning and professional training in specialized fields such as veterinary medicine. Finally, a set of implications of design for future systems is provided.

ACKNOWLEDGEMENTS

I would like to express gratitude to Francisco Ortega for his guidance, leadership, and support throughout the course of my research. I would also like to thank Alyssia Miller De Rutté for her expertise in Spanish studies and her insightful guidance. My sincere thanks go to Nikhil Krishnaswamy for his continued support and advice. I would also like to thank the NUI lab and Computer Vision labs, and all of my peers in the Computer Science Department for their friendship, support and knowledge across various fields. Specifically, I would like to thank Phat Ho for his huge role in creating the Unity Environment that we originally were going to use for this experiment, as well as finding and reading related works. I would also like to thank Richard Rodriguez for his guidance on Thematic Analysis, specifically how to cluster the results and create a graph. And I would like to thank Swagatalaxmi Aich for her help with brainstorming tags for the transcripts. I could not have done this without the emotional and academic support from my friends and family.

DEDICATION

I dedicate this thesis to my friends and family, whose unwavering support and encouragement have inspired me to pursue both computer science and foreign languages. Thank you for believing in me!

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1

Introduction

Spanish is the official language of 21 countries and is widely used in many more [1]. In the United States (US), more than 67.8 million people speak a language other than English at home, and Spanish is the native language of 42 million people [2]. Due to the number of Spanish-speaking clients worldwide, there is a growing need for multilingual veterinarians. Medical care depends on clear communication between the healthcare provider and the client [3]. Although many native Spanish speakers also know English, it can be difficult to understand medical terms in a second language. People might not be familiar with highly technical terms in their native language and are even less likely to be familiar with them in a foreign language. This is why it is important to provide people with medical and veterinary care in their native language if possible.

In medical and veterinary medicine education, students must be well prepared for the diverse situations they may encounter in practice. Learning how to interact with clients is a key component of any veterinary program, and standardized clients are commonly used to provide this experience [4]. Traditionally, this includes students interacting with an actor who will play the role of a pet owner, while the student conducts a mock consultation [5]. This method has limitations, requiring actors to not only participate, but also learn and memorize the pet's symptoms and backstory. Current technology allows us to simulate student-client interactions in a virtual environment. With this virtual tool accessible on laptops and tablets, students can practice these interactions anytime, anywhere, without relying on another person's availability or time commitment.

Virtual environments have become increasingly helpful tools in education and training [6]. They allow learners to place themselves in simulated situations to practice skills essential to their fields of study and grow their confidence in themselves, before applying those skills to a real-world setting. Often virtual environments can provide students with opportunities to access situations they would not have otherwise. This is very useful for students trying to learn skills they may not have the opportunity to practice without technology.

Previous work has shown that virtual reality (VR) improves the learning process in many subjects, such as science, mathematics, and music [6]. Virtual environments have also been shown to help with language acquisition. Previous work has found that VR increases engagement and appeal in language learning settings [7]. Although there has been some research on language learning in VR, there are still many areas to explore the impacts VR can have on making language learning more effective. VR has also been used for exposure therapy to help people overcome specific anxieties, phobias, [8] trauma, and post-traumatic stress disorder (PTSD) [9, 10]. VR has been used to study attention deficit disorders [11]. Although many studies focus on VR, virtual environments based on desktops have also been shown to influence user embodiment [12].

This paper explores the user experience of MeTabi, an application that uses a virtual environment designed to learn a second language through task-based interaction with generative artificial intelligence (AI). Participants completed a virtual consultation in the application, sharing their likes, dislikes, and suggestions for improving the user experience, and participated in a semi-structured interview at the end.

A qualitative thematic analysis [13] was performed on participant data to identify recurring comment clusters. Based on the findings and common participant criticisms, a set of design implications for future systems was developed.

Research Questions:

RQ1: What are the common themes in the participants' feedback on learning veterinary Spanish through this virtual environment?

RQ2: How engaged and immersed do participants feel while interacting with the virtual environment?

RQ3: What specific features did the participants find useful or challenging and what suggestions for improvement were offered?

Chapter 2

Related Work

Immersion has consistently been shown to be one of the most effective methods for learning a second language [14, 15]. Learners achieve significantly better results when placed in an environment surrounded by the target language and are required to actively use it. In contrast, when learners are removed from such immersive settings, their language skills often decline [16, 17]. In addition, studies have found that interactivity plays a crucial role in enhancing learning outcomes [16, 18].

2.1 Task-based language learning

A well-established approach to language education is task-based learning (TBL), prioritizing practical tasks as a key pathway for developing language skills. Immersion has long been shown to be an effective method of learning a second language [15, 19]. By focusing on communication and problem solving in immersive real-life contexts, TBL has proven effective in enhancing speaking, listening, and comprehension abilities through meaningful engagement in the target language [20, 21].

2.2 Language Learning in Virtual Environments

The use of virtual environments in language learning has begun to receive attention in recent years. Research has shown that immersive, task-based virtual simulations can enhance language acquisition and professional skills, particularly in fields requiring specialized communication. For example, Petersen [16] explored the impact of VR on language acquisition, finding that VR improves situational interest and embodied learning.

Studies have shown that VR can offer a more immersive and enjoyable experience compared to traditional language learning applications, regardless of the preferred styles of the learners [7, 22, 23]. VR has been found to increase the motivation of learners [24], often through interactive tasks

where users can manipulate objects and learn their names in the target language [22]. Another common VR task involves search and find activities, in which participants receive instructions in the target language and must locate objects within the virtual environment [22]. Language learning in an intermediate Spanish VR medical setting showed that extended use led to users feeling more confident in seeing themselves using Spanish in a future work environment [25].

One of VR's greatest advantages is its ability to provide immediate feedback on linguistic input, which is crucial for language acquisition and vocabulary development. The research of Miller and Gildea [26] highlights that children learn new words most effectively through interactive methods, such as conversations and writing. They emphasize that ideal learning occurs when an automated system offers real-time explanations of unfamiliar words in context, mimicking the natural inquiry process of live conversation [6].

Language learning in augmented reality (AR) has also shown improvements in vocabulary acquisition, as well as reading, speaking, and writing skills [27, 28]. These examples illustrate the transformative potential of VR and AR in facilitating language comprehension and acquisition.

2.3 Virtual Patients for Medical Settings

Using virtual patients in and out of a fully virtual environment has been an increasingly popular tool used by medical teachers. To ensure that students have access to various experiences treating different cases, standardized patients (SP) have become commonplace in medical school [5]. In 1971, Harless et al. described the use of a computer-aided simulation of a clinical encounter [29], providing students with the opportunity to simulate real interactions with patients without the stress or liability of affecting the life of a real person [5]. Virtual humans have also been used in settings such as helping college students with mental health [30]. The use of virtual human interactions has also been shown to increase empathic interactions with suicidal patients in clinicians [31].

Chapter 3

Methodology

Figure 3.1: A screenshot of the Unity Environment we developed. The participant can practice asking a question to the client, and then click on one of the buttons in the Menu box. After pressing the button a pre-recorded response from the client is played for the user.

3.1 Unity Application Development

I collaborated with my research partner Phat Ho to create a virtual medical environment that allows users to simulate the role of a medical assistant or nurse in a Spanish-speaking hospital setting. We created the environment of a medical consultation room using Unity, with a realistic-looking avatar made using Character Creator v4.4. In this environment, the user is given the task of collecting information from the client, such as their name, date of birth, and information about the symptoms they are experiencing. The user can click on speaker buttons within the environment to hear pre-recorded responses from the virtual client. This was developed to be used in a VR headset, we tested it using a Meta Quest 2. Although we spent a lot of time developing this environment, we ended up using MeTabi – an existing application that we were not involved with the creation

of – for the current research study. We would like to use the Unity application that we created for a future research experiment.

3.2 Current Research Experiment

Although the initial environment was set aside for the current study, my focus changed to conducting a Qualitative Thematic Analysis of the MeTabi application. The following sections describe the MeTabi application, participant recruitment, experimental procedures, and analysis of the results.

3.3 Participants

For this study, eight participants (two male, six female) were recruited to participate from a Veterinary Spanish class at Colorado State University (CSU). Six of the participants were students and two were instructors. Their ages ranged from 24 to 34 years old (M = 27.1, SD = 3.1). Any student or instructor involved in the class was considered eligible for the study. Because of this, all students had similar levels of knowledge in both Spanish and veterinary sciences. Both instructors were native Spanish speakers, offering a different perspective on the application as opposed to the students who were in the process of learning Spanish.

3.4 Application

For this study, we used the MeTabi website, which allows users to access lessons and virtual environments designed to facilitate language learning through interactive activities that combine language and technical skills for an encompassing experience. MeTabi used OpenAI's GPT-3.5 model to provide users with responses from virtual people within the environment. This application was used for homework assignments in the class, so all students and instructors were familiar with the application before participating in the study because the experiment was run about 3/4 of the way through the semester. Figure 3.2 shows the main page of MeTabi, where users choose which module they wish to work on after logging in.

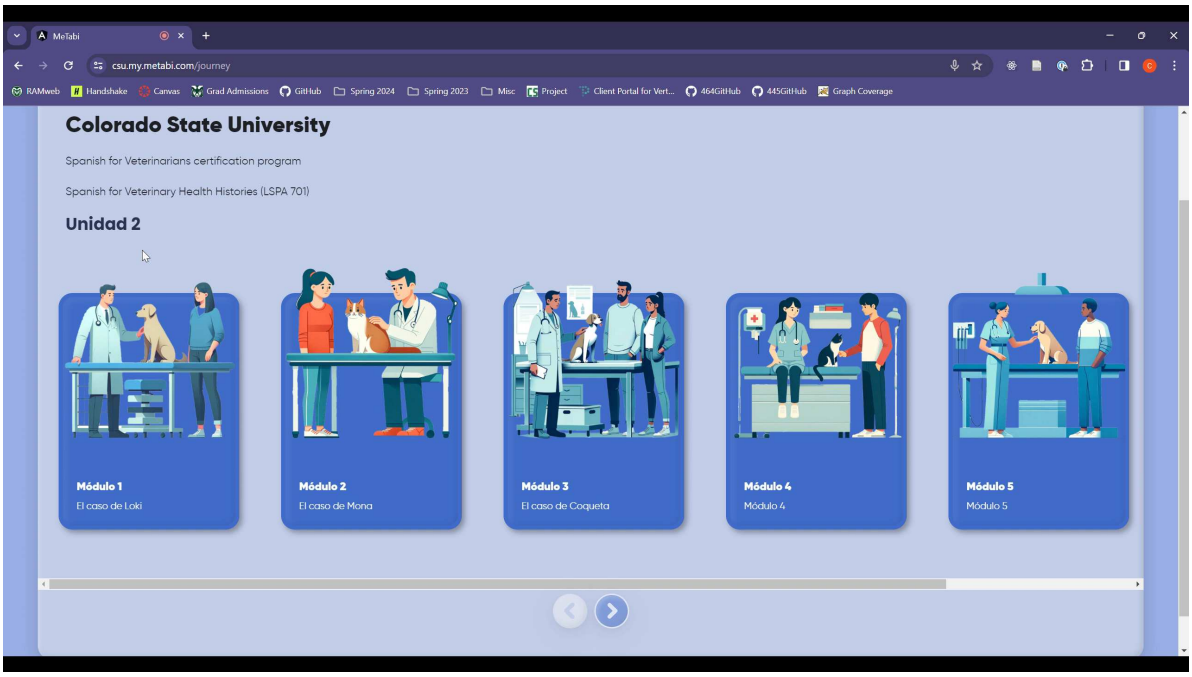

Figure 3.2: A screenshot from a participant video. The participant is on the Módulos page of MeTabi, the landing page when logging into the application. Participants were given the option to choose between Módulo 1: El caso de Loki and Módulo 2: El caso de Mona for the experiment.

In this study, students were told to complete a "consulta virtual" (virtual consultation). They were given the option to choose either the consultation from "módulo 1" (module 1) or "módulo 2" (module 2). Each module had 3 different clients from which participants could choose. Most of the participants chose the client closest to the entrance. These modules were chosen because they covered topics the students had already covered in class. Figure 3.3 shows the virtual environment during a virtual consultation with a client.

3.5 Equipment

All participants utilized the same Alienware x15 R2 laptop for the experiment, rather than their personal devices. The laptop operated on Windows 11 Pro and featured a 12th Gen Intel® Core™ i7-12700H processor with a clock speed of 2.30 GHz. The version of MeTabi accessed during the experiment was the version online between April 2nd and April 26th, when the experiment was conducted.

3.6 Procedure

The study consisted of three main steps after signing the consent form: completing a virtual think-aloud style consultation, a semi-structured interview, and the demographics survey.

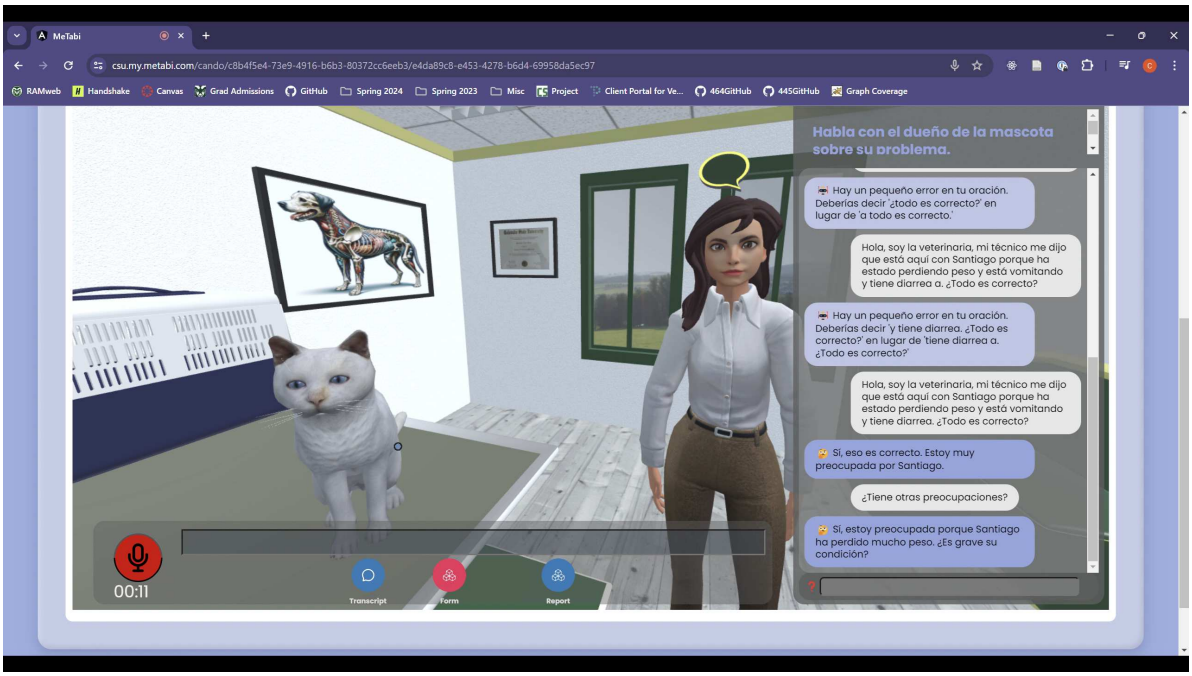

Figure 3.3: A moment captured from a participant video. The participant is seen engaging in a consultation with a virtual client about her cat. The transcript from the conversation is shown on the right side of the screen, with the AI's responses left justified and the participant's on the right.

3.6.1 Virtual Consultation

Each participant was instructed to complete a virtual consultation activity in MeTabi while employing the think-out-loud method [32] while completing the consultation. While they were completing the task the researcher would encourage them to speak their thoughts aloud periodically or ask them prompting questions on which aspects of the application they like or dislike based on their interactions with the application. Throughout this process, the participants' interactions were screen-recorded along with audio of their conversations with the researcher.

3.6.2 Semi-Structured Interview

Following the consultation, a semi-structured interview was conducted with the participants where they were asked more about their likes, dislikes, experience, and ideas for improvements within the application. The survey incorporated selected questions from the short-form user engagement scale [33] to assess participants' perceived engagement with the lesson and learning task, along with a few additional questions regarding their willingness to continue using the application.

3.6.3 Demographics Survey

After the interview, the participants completed a survey asking for demographic information, including gender, age, race, and spoken primary language. After the survey was completed, the participants were compensated with a $20 Amazon gift card.

3.7 Data Analysis

I conducted a qualitative analysis of all the participant recordings. I began with the transcription of each video, utilizing Microsoft Word's transcription service to generate initial drafts. I then meticulously reviewed each transcript to ensure accuracy, leveraging my familiarity with the experiments and conversations to clarify unclear speech. Due to the limitations of the transcription software, which could only transcribe one language at a time, I typically set it to English and manually transcribed the Spanish audio. This required substantial editing.

Once all of the transcriptions were completed, I conducted a thematic analysis. I read each transcript to identify emerging themes, annotating each line with the labels: aesthetic, app features, bugs, interaction preferences, lesson, negative feedback, positive features, and proposed features. I confirmed these labels with my advisors to ensure the relevance of my tagging. After this, I compiled an Excel sheet documenting the common criticisms raised by participants during their participation in the study, organized by app feature, and tracked which participants mentioned each point.

Finally, I grouped all of the criticisms into clusters. Throughout the thematic analysis and clustering process, I discussed and validated my findings, asked for advice, and confirmed the themes with my advisors and colleagues in my advisor's research group to ensure the precision and relevance of my tagging and clustering.

Chapter 4

Results

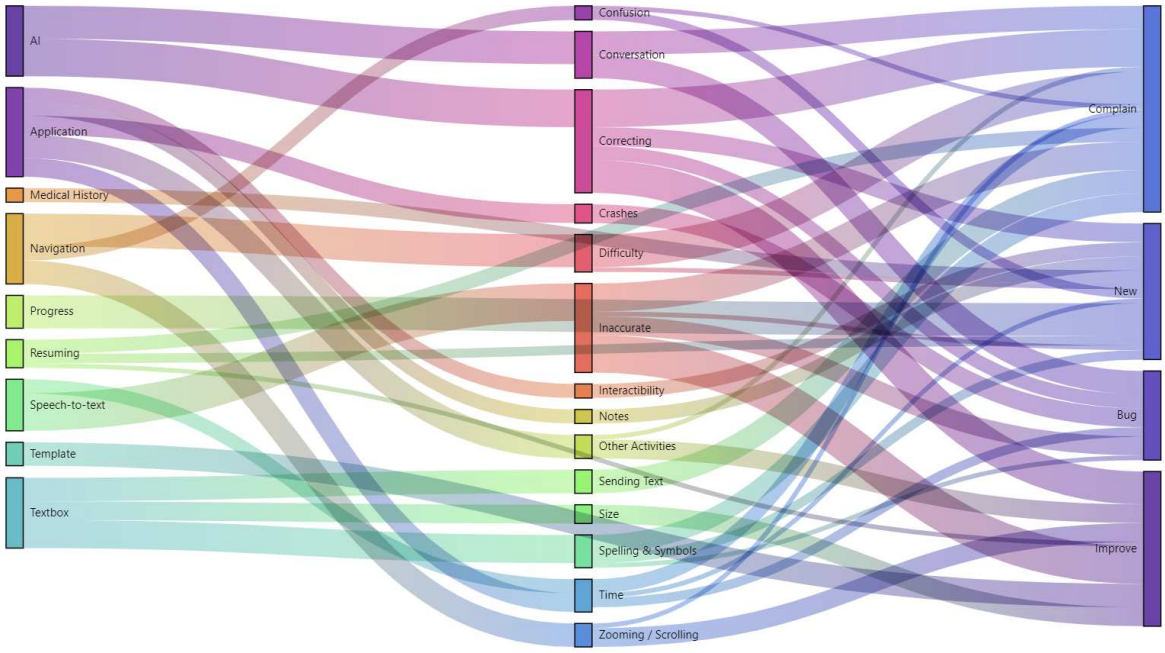

Figure 4.1: This figure illustrates the process of theme development in the analysis. On the left, it lists the features that were criticized by participants. In the center, the emerging themes of these criticisms are displayed. On the right, each criticism is categorized based on whether participants proposed a solution and the type of solution suggested. The thickness of the connecting lines reflects the number of participants who commented within each category.

4.1 Features

The participants provided critiques of the features in the virtual environment and provided various suggestions for improvement, as seen in Figure 4.1. In this section, I will highlight the most common criticisms and proposed enhancements for each feature. Figure 4.2 details all of the features visible during the virtual consultation. The features are presented in order according

Figure 4.2: Features on screen during a virtual consultation: Speech-to-Text is accessed using the red microphone button in the lower left corner of the screen. After the text is transcribed from the user's speech it appears in the Text Box underneath the virtual client. The AI responds in the chat box transcription on the right side of the screen, with AI responses justified to the left, and user responses indented to the right. The Form (Plantilla de Tareas) is accessed by clicking on the "Form" button at the bottom of the screen under the Text Box, after which it will pop up in the middle of the screen on top of the virtual client.

to the number of participants who critiqued them, from most to least, as seen in Table 4.1. In this section, I only cover features brought up by at least half of the participants (four or more). The Other Activities category refers to comments participants made about activities other than the virtual consultation, and thus are not discussed in this paper, since the focus of this paper is on the virtual consultation.

4.1.1 Speech-to-Text

The speech-to-text was the most heavily criticized aspect of the application. The majority of complaints came from the inaccuracies of the transcriptions, whether that was putting something other than what the user said, not using proper grammar, or sometimes not picking up what had been said by the participant at all.

Inaccurate Transcriptions of Participants' Speech

The inaccurate transcriptions of participants' speech were a common source of frustration in the application, especially with how picky the program was in making sure that participants said

Table 4.1: This table shows the clusters for each feature, derived from the thematic analysis. Each cluster is listed by the number of participants who made comments that were categorized in the corresponding cluster.

Feature	Cluster	# of Participants
AI	Correcting	8
AI	Conversation	8
Speech-to-Text	Inaccurate	8
Navigation	Difficulty	7
Progress	Progress	7
Text Box	Spelling & Symbols	6
Other Activities	Other Activities	5
Text Box	Sending Text	5
Form	Template	5
Navigation	Zooming / Scrolling	5
Time	Time	4
Environment	Crashes	4
Resuming	Resuming	4
Text Box	Size	4
Medical History	Medical History	4
Environment	Interactability	4
Notes	Notes	3
Speech-to-Text	Loading Time	3
Form	Feedback	3
Environment	Avatar	2
Application	Continued Use	2
AI	Voice	2
Environment	Aesthetic	1

things with perfect spelling and grammar (discussed in 4.1.2). The speech-to-text often wouldn't pick up what the user had said, sometimes putting completely different Spanish words, sometimes it would even transcribe English words when they had been speaking in Spanish, and oftentimes not transcribing anything at all, or putting "undefined".

Another common complaint was that the speech-to-text system would often stop transcribing at a certain point and only transcribe the first half of what was said. This would often happen when the user took a brief pause speaking either in between sentences or while trying to think of a word. One participant expressed interest in a button that would let you pause the recording while trying to think of a word or phrase, but most participants simply wanted the text-to-speech to continue transcribing regardless of when you paused briefly in the middle of speaking.

The text-to-speech especially struggled with picking up names, usually when the user referenced the name of the pet they were trying to diagnose. Stumbling over a word or stuttering especially resulted in the text-to-speech transcribing gibberish. One participant said that they wished the speech-to-text would ignore when they used ums and uhs while speaking, because it would try to interpret those as words.

Loading Time

Three of the participants made comments about the length of time it takes to generate the transcriptions. Participant 1 mentioned it being annoying when they know that they said something wrong and they still have to wait for the transcription to be generated, saying that they wished that there was an option to cancel the speech-to-text so that they could just rerecord what they meant to say correctly. "I guess the loading, too, takes kind of a while, like when you mess up. [...] I wish it, instead of pausing and starting the microphone, like if it started and you mess up, you could just put cancel and it won't read anything. That'd be nice."

4.1.2 AI

There were three main points of contention when it came to the interaction with the AI. These included the conversation, the way the AI corrected the student's speech, and the voices of the AI patients.

Correcting

One of the most common criticisms of the application was over the way the AI agent dealt with errors the participants made in their speech. Although there were differing ideas as to what the agent should do in response to mistakes, all eight of the participants felt that the AI was too strict in it's correction of their mistakes. The AI would stop the conversation every time a grammar, spelling, punctuation, typographical, or capitalization error was made, as well as if the AI assessed that the user had said something irrelevant or unprofessional. The agent would not continue the conversation until the user submitted a sentence with perfect spelling and grammar. For example,

one of the most pedantic corrections made by the AI agent was when there was an extra space between two words. "I mean yes [it's annoying], because it's not that the meaning is not getting through. It's just correcting like there's an extra space" said Participant 7.

Participants gave many alternate solutions to this complaint. Seven of the participants said they wanted the agent to continue the conversation even after a mistake was made. Five participants recommended that the program could simply point out their mistake and/or how their sentence could have been said correctly, but then continue the conversation by still responding to what they said. A couple of participants thought it would be nice if the program gave them one chance to correct their mistake, and then move along with the conversation. Some participants said that the program should not even be correcting them on certain things, such as in the example of the extra space between words, incorrect capitalization of words, simple grammar mistakes, or lacking an accent mark on a word. Many students pointed out that the virtual consultation was supposed to mimic a real-world conversation with a client, in which case capitalization and spelling would be irrelevant because they are speaking. This issue was also related to the student's criticism of the speech-to-text, noting that often the things that the AI was correcting them on were products of the transcription of the speech-to-text and not an issue with what they had spoken. However, other participants expressed that they appreciated getting feedback on their mistakes, but simply did not like having to re-dictate or re-type their sentences multiple times to continue with the conversation.

Another problem was the program telling users that something they had said was irrelevant or unprofessional. These were tricky because they would happen sort of randomly and leave the user wondering what it was that the program thought was inappropriate about what they had said. This seems to be a feature implemented to keep students on topic and dissuade them from fooling around with the program.

Participant 3 shared two examples where the AI claimed their input was irrelevant. In one instance, they asked about flea prevention after observing fleas on the virtual pet, but the AI client responded "That's not relevant." The participant explained, "I was like, 'well, it's kind of relevant [...] you see fleas, are they on prevention?'" In another case, the participant inquired about the type

of food being given when a pet was not eating, to which the AI similarly replied, "That's not relevant." The participant felt these were important questions directly related to the pet's symptoms.

Conversations with the AI Agents

A common complaint about the AI conversation is that the AI would occasionally contradict itself or give false information to the user. An example of this is when Participant 3 commented "My vet tech was masculine when he was out in the lobby, but now he's saying that it should be female."

The AI would also sometimes argue with the user over things that the users did not think were worth fighting over. An example of this is when Participant 2 told the AI "Veo que hay mocos verdes en su nariz." (I see that there is green mucus in her nose.) To which, the AI agent responded, "No, no son verdes. La secreción nasal de Lulu es clara." (No, it's not green. Lulu's nasal secretion is clear.) The participant reacted with "Really? Alright well sure. I would call that green, but that's fine, sure. Like, I can see [the mucus]." But rather than argue with the AI, the participant just responded with "¿Hay secreciones nasales claras, verdad?" (Her nasal secretions are clear, right?)

Some other comments made by participants were that sometimes the avatar would not respond after saying something to it, leading the participants to just copy and paste their message and send it again so that the avatar would respond.

Other complaints were regarding the beginning and ending of conversations. Each time a user approaches a virtual person in the environment, the conversation always begins with the same opening line. This can be annoying if you leave a client's "area" in the middle of trying to address their pet's problem and re-enter the area. For example; you might have just been talking to the client about their dog's mucus and you step away to look at the dog's nose and step back in front of the client and they immediately say "Hola." (Hello.) It makes it seem like the client forgot they were talking to you and are trying to start a conversation from scratch. It leads to confusion from the user's perspective because they do not know if the model remembers the previous conversation with the user.

Participants also mentioned how it is often awkward trying to leave the client because there is no clear end to the conversation. The agent will continue to respond as long as you keep talking to it. A couple of participants said it would be nice if the AI somehow indicated when it was satisfied with the conversation and ready to end the interaction.

Voice

Two participants had criticisms of the voice of the avatars. One participant expressed that the AI voices were too robotic and didn't show enough emotion, so it was hard to take them seriously when they talked about their concern for their pets. Another participant pointed out that all of the people you interact with in the facility have the same Spanish accent, and it would be a helpful addition to the program to include a more diverse set of Spanish accents, to both make the experience more realistic and provide users the opportunity to learn and understand the Spanish used by people from different regions they may encounter worldwide.

4.1.3 Navigation & Controls

Difficulty Moving Through the Environment

Seven of the eight participants made comments about navigating through the virtual environment being difficult. Notably, three of these participants said that they did not play video games, and that might be why they found the environment more difficult to navigate. Some students mentioned that specific methods of controlling the navigation were more challenging than others. Two participants said navigation with the mouse pad was more difficult, while one said using the mouse was more difficult.

Three of the participants still didn't know about some of the options for navigating the environment after using the program for over a month already. I told two of the participants that they could use WASD to move instead of the arrow buttons and they hadn't realized before. Participant 8 discovered that when using a mouse to navigate you can teleport, which I hadn't realized before then either.

Some of the proposed solutions for these criticisms were adding instructions or a tutorial when you try a virtual consultation for the first time. Another participant recommended a menu or button including a list or tutorial of the controls available to navigate the environment. Participant 6 recommended having buttons like the ones in Google Street View to move you up and down a road to make moving easier and said that they could even include buttons that take you straight to the clients' rooms instead of needing to navigate freestyle. Participant 1 also said that navigating the virtual clinic was cool but not necessary to the goal of the application: practicing Veterinary Spanish.

Zooming & Scrolling

A specific bug with navigation in the application was that when using the arrow keys to move through the environment, the webpage scrolls up and down. This is annoying to users because they have to scroll the page back down to view the environment constantly. When trying to navigate the environment on a touch screen, you swipe with your finger, which leads to the webpage zooming in and out, and if you swipe your finger across the screen in a specific way, it can lead to the browser taking you to the previous page, consequentially causing the user to lose their current progress in the environment. These bugs should be fixed to make navigating the application an easier experience. They can lock the arrow function to move the virtual environment without scrolling the webpage.

4.1.4 Progress Through the Modules (Módulos)

Another common criticism of the application is that there is currently no way to track your progress through the modules and activities. Seven of the eight participants brought up this criticism and all of them offered up suggestions for ways to better keep track of your progress through the application. Six participants suggested adding a checkmark next to the modules and activities the user has completed as an easy way to keep track of what has been done. Two participants recommended adding an area that could list the veterinary topics you have covered in the application so far. Two participants recommended adding a progress bar either in the upper right corner

or the bottom of the screen to keep track of how many activities through a module you currently have completed, and how many you have left. Some examples of this were stars symbolizing the activities, a standard progress bar, numbers such as 12/15, or points for completing activities.

Participant 2 recommended highlighting the next module or activity you have to complete to make it easier to know which one to work on upon returning to the app, and that it would be nice to be able to play back recordings of your responses from previous activities to hear how your pronunciation had improved over time. Participant 3 offered the option of a self-test users could take to see if they increase their score over time. Participant 8 was excited to recommend a feature giving the user an award for each pet they see, and the award would include a picture of the pet's face with their name.

4.1.5 Text box

The text box was where the users' dictated text would appear after being processed by the speech-to-text. Users were allowed to edit the text in the text box before sending it to the AI model, to correct spelling or grammar mistakes made by the speech-to-text. The user could also start typing in the text box without having to dictate the text through the microphone. This gave users the option of simply typing their conversations with the AI without having to use the speech-to-text feature. Previously, participants did not have the option to edit the text generated from their speech before it was sent to the AI, so they expressed great appreciation for this newly added feature. However, despite the improvement, several concerns were raised about the implementation of the text box, which I explore below.

Spelling & Symbols

Five of the participants struggled with typing language-specific characters in the text box. These are characters that are used in the Spanish language that are not used in the English language, such as the letter ñ, accent marks or diaeresis over vowels (á, é, í, ó, ú, ü), and the inverted question and exclamation marks (¿, ¡). Since these students are studying in the United States, they are using American English keyboards without the ability to type these characters. Many comput-

ers have the option to switch your keyboard layout for different languages, which, unless they are using a touch screen device, does not change the physical keyboard, only what the keys on their keyboard are mapped to, so users must remember which English key is mapped to which Spanish key. Setting this up takes time and knowledge of your system's settings. Not everyone knows how to change their keyboard settings in this way.

A common strategy for typing language-specific characters among participants was to copy-paste them from the chat history if the AI had used it previously. In many cases this meant typing a message without using the correct characters, waiting for the AI to give them the corrected sentence, then pasting that into the text box to re-send to the AI. This solution requires the user to send every sentence with special characters twice, making the conversation much longer.

Participant 6 expressed interest in having buttons in the MeTabi application that users could press to insert a language-specific character into the text box without needing to do anything fancy with their computer settings. Some other language learning applications provide this. Participant 2 expressed interest in spell check in the application, saying "Sometimes I will like, before I put what I'm going to write in there and click submit or whatever, I will run it through Google Docs cause it has got pretty good Spanish spell checking."

Submitting the Text from the Text Box to the AI

Text is submitted to the AI by pressing enter after making any edits to the text the user wishes to change. Two participants had experiences where they didn't have the text box selected (meaning their last click wasn't inside the text box, but somewhere else in the application) so when they pressed enter, the microphone started to record. This puts them in a situation where they must click inside the text box again and press enter to send their message or else all of their current text will be lost and replaced with whatever the microphone and speech-to-text picks up in the meantime.

Another confusing part of the application that two participants pointed out is that there are two text boxes in the application, one of which sends messages to the avatars in the environment, and another one that is somewhat unclear but seems to be a way to ask for help using the application.

However, the second text box is right underneath the chat transcript with the avatars, so it is very misleading and looks like it should be the text box for interacting with them.

The Text Box Size

Four of the participants said that the text box provided is too small. Users cannot view the entire text when speaking in full or multiple sentences and must scroll through it using the arrow keys. Every time the user clicks outside of the text box, it automatically resets to the beginning of the text, meaning if they want to reference or edit something or add a new sentence, they must move the cursor through the text every time. The solution to this was simple. The participants said that the application should either have a larger, paragraph-style text box or let the text box expand based on the size of the input from the user.

4.1.6 Form (Plantilla de Tareas)

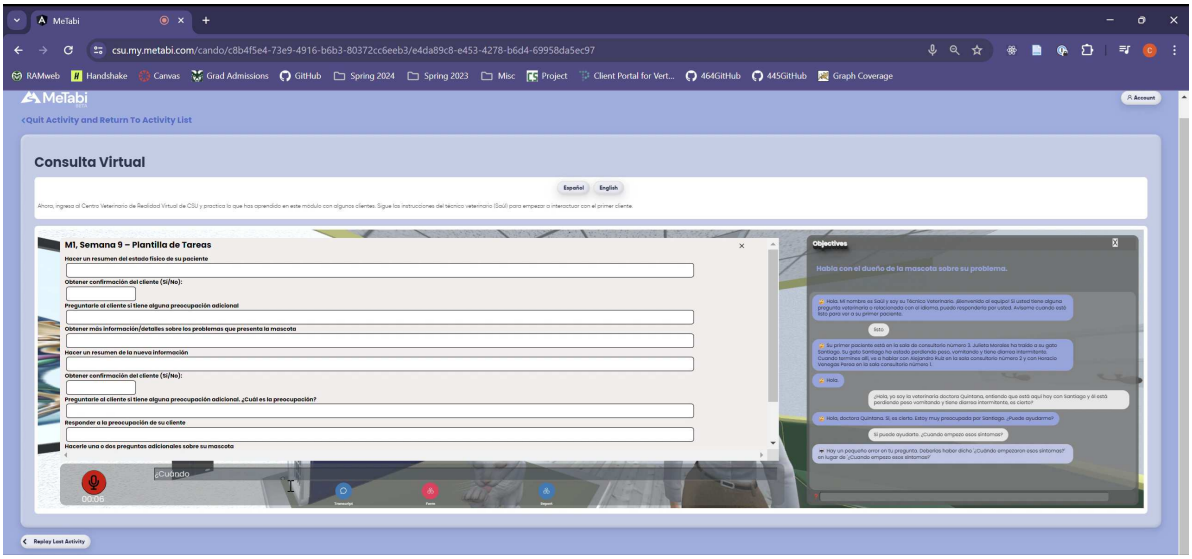

Figure 4.3: The Form (Plantilla de Tareas) appears on the screen whenever the "Form" button on the bottom of the screen is pressed, if the button is pressed again, the form disappears to show the virtual environment behind it again.

Five of the eight participants said that the order of the form (Plantilla de Tareas) did not make sense or flow well with the natural conversation they were having with the client and wished it was

more organized, shown in Figure 4.3. A common complaint was that the form was repetitive, and the user was told to summarize what the client had said after most prompts. It was also long and wordy. It was almost as if it was trying to be a script to follow to talk to your client, but it was simply too wordy. It would repeatedly tell you to ask the client if they had additional concerns, tell you to summarize what the client had just said, and then ask the client for confirmation that your summary was correct. Many participants complained that this should be reserved for the end of the conversation, not done repeatedly, and it could all be combined into one step: summarize and confirm with the client what they told you. Some participants thought the template would be more helpful as just a bullet-point list of topics to address with the client instead of laying out the whole conversation step by step for you. One participant had a different opinion, stating that they thought the conversation was too open-ended and weren't sure what to say, stating that they wished the conversation was laid out for them more the first time using the application. One participant wished that the form was based on a real veterinary clinic form like what they will receive once they start performing real consultations.

4.1.7 Lesson Duration

Four of the eight participants stated that the activities and modules took a lot of time to complete. Two participants mentioned that often the virtual consultation takes a long time to load, and sometimes doesn't load at all and must be rebooted. There were also complaints that the application took a lot of computing power from the computer, causing laptops to heat up and run the fan, lagging, draining the battery quickly, and occasional crashes. One participant said that learning how to get around all of the bugs in the program causes the program to take longer than it should. One of the participants who was an instructor and native Spanish speaker noted that as a native speaker, they should be able to breeze through the virtual consultation, but instead, they were not able to because they had to deal with the speech-to-text not transcribing them correctly, and the AI making pedantic corrections to their sentences.

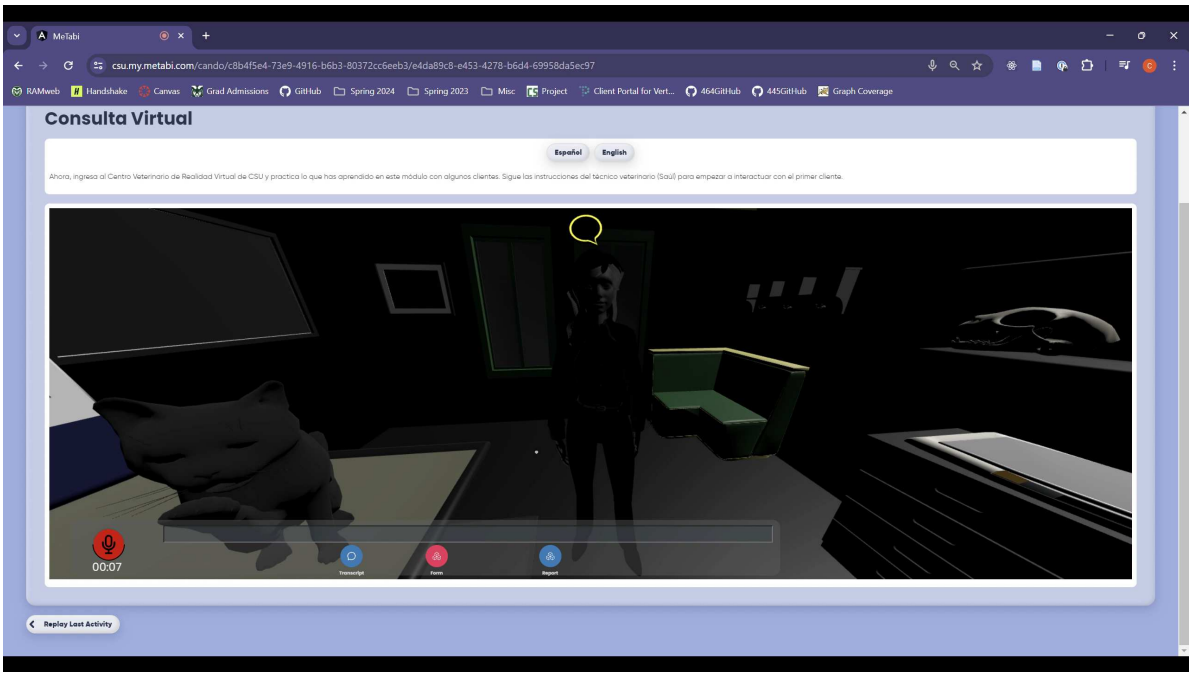

Figure 4.4: Screenshot from a participant video, when the environment glitched in the middle of the consultation and the screen went dark.

4.1.8 The Virtual Environment

Crashes

Four participants mentioned that the application would occasionally crash completely on their personal computers. Three also described a recurring bug in which the virtual environment would suddenly turn black as if all the lights had gone out. During one of the recorded studies, this issue occurred, Figure 4.4 shows a screenshot of this glitch. The participants noted that it was usually followed by the entire application crashing shortly afterward.

Interactability

Three participants expressed interest in more interactability in the virtual environment. Participant 1 was confused about whether there was a way to interact with the objects in the room, and after trying and not figuring out a way to, expressed interest in being able to do things such as take the pet's weight and temperature during the consultation, as well as interact with the other objects in the room. "All the little details, like it's cool and all, but what's the point if you can't use them?

The thermometer and the scale and stuff like that. Cause that's another thing that we've been learning in our curriculum is how to say temperatures in Spanish or how to collect weight in Spanish and stuff, so if I can't do it, it's like, how am I gonna interplay that stuff into this program?"

Participant 1 was not the only one who felt this way. Participant 6 also said "I feel like this interaction isn't finished because I would then want to do an exam with them. I'd want to do diagnostics. I would want to talk about the results and be like, this is what I actually think it is." Participant 8 said that they wish the environment were a fully immersive VR experience. "I can't wait to use the goggles and go actually inside the whole clinic and talk to the avatars in person. That would change the whole experience, I'm waiting for that."

Pets

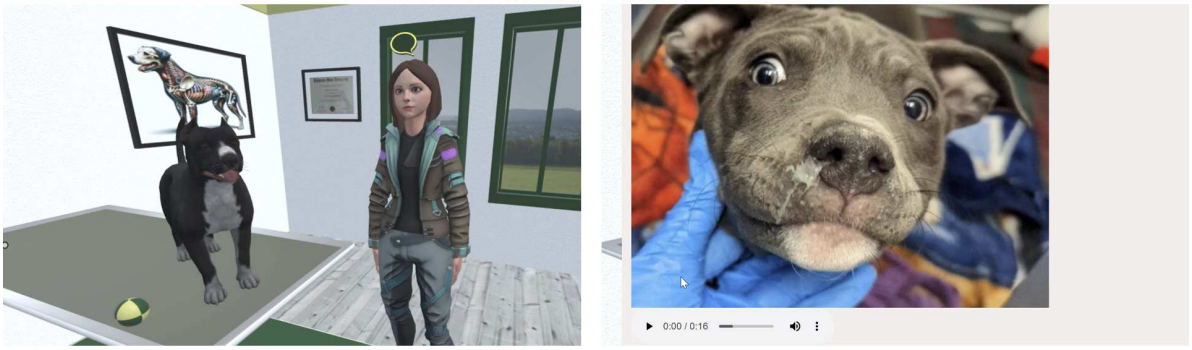

Figure 4.5: The left image shows a hot spot in the virtual invironment, the yellow and green ball on the table next to the pet. The image on the left shows the pop-up that appears after the hot spot is clicked on, in this example there is a picture of a dog with mucus in it's nose, and an audio clip of the dog coughing.

While the majority of participants enjoyed the virtual pets with participants describing them as "cute" and "fun," a couple of participants expressed confusion or dissatisfaction with the interactions with the pets. Interaction with the pets was facilitated through virtual hot spots. These hot spots were virtual markers that appeared on or next to the pet, allowing users to click and access media related to the pet's treatment, as seen in Figure 4.5. For example, clicking the hot spot might display an image of mucus coming out of the pet's nose, or audio of the pet coughing. Participant 5 stated that they did not realize what the hot spots on the pets were at the beginning, or how to

interact with them. They had multiple ideas for reducing the confusion; a menu or tutorial explaining controls, a pop-up label saying "click here" over the hot spot, or replacing the hot spot with a button at the bottom of the screen next to the Form. Participant 8 did not like the hot spots, saying they wished that the information was available on the virtual pet. For example, including the mucus on the nose of the virtual pet, instead needing to click to see a picture of the mucus.

4.1.9 Resuming

Three of the eight participants expressed the desire to have the option to exit and then resume their session in the virtual consultation. They found that having such flexibility would allow them to better manage their time and address unforeseen interruptions during consultations. This feature would enable participants to maintain continuity in their sessions, benefiting from a more personalized and accommodating service without losing their session progress or context.

The option of resuming a session was added to the application at some point, but there was a lot of confusion surrounding this feature. Some participants were not aware that they could resume a session, or did not trust that the system stored their previous conversation, because nothing in the application indicated that it had, but instead started the user back at the beginning of the consultation with no chat transcripts pulled up. One participant said that it would be good if the application placed you back in the position you were in before with your old conversation still in the chat transcript, to make it apparent that it had been saved. Otherwise, there was no indication that anything was saved and most participants would restart from the beginning to play it safe.

4.1.10 Medical History (Historia Médica)

Before starting consultations, a virtual veterinary technician (vet tech) lists all the pets waiting to be seen and which room they are in, in addition to describing the symptoms each pet is experiencing. This information is also posted outside each consultation room, in a form titled Historia Médica (Medical History), shown in Figure 4.6. Two participants mentioned that the monologue was too long and overloaded with information, making it overwhelming and difficult to remember everything. They recommended that the vet tech only describe one patient to the vet student at a

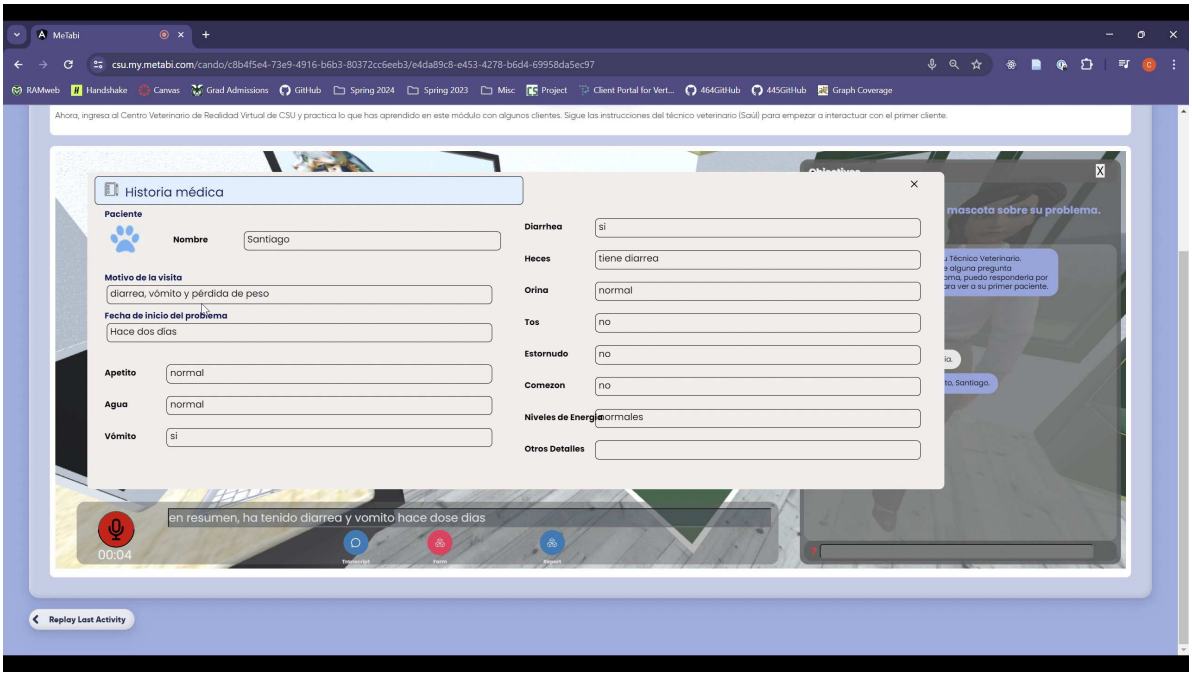

Figure 4.6: The Medical History (Historia Médica) appears on the screen whenever the clipboard outside of a consulation room is pressed, and in some of the rooms there is a computer on the counter that can also be pressed to pull up the medical history. If the X in the corner is pressed, the medical history disappears to show the virtual environment behind it again.

time or that the introduction from the vet tech could be skipped altogether because of the medical history posted outside each client's room. Three participants mentioned that they wished the medical history forms were viewable from inside the consultation rooms. A proposed way to do this was to include a button at the bottom of the screen the same way the template form is accessed.

4.2 Experiences

Although participants had numerous complaints and suggestions for improving the application, the overall experience was that it provided a unique and engaging alternative to traditional homework methods, such as recording videos of a one-sided conversation. Despite participants spending 62% of the study voicing criticisms and only 38% making positive comments, this is likely due to the application being in beta and still experiencing bugs, which motivated participants to focus on issues as they arose. However, all agreed that the application was more immersive and engaging than the conventional homework and study approaches.

4.2.1 Engagement and Immersion

When asked if they found the application engaging, participants generally felt that the interactive nature kept them engaged, although it could become frustrating when the application dragged out conversations by repeatedly correcting them. Similarly, most participants said they felt immersed in the experience, with real-time conversations with the virtual client being the primary factor contributing to the sense of immersion. Several participants mentioned that it looked like they were inside the CSU Veterinary Teaching Hospital (VTH), which helped them to feel more immersed. However, frustrations with the application's pace and the need to retype everything were cited as factors that broke immersion for some users. Although most of the participants appreciated the freedom to interact in a more realistic and open-ended environment, only one participant preferred more structured guidance, which they felt could enhance their engagement and understanding.

4.2.2 Aesthetic Appeal

The participants found the environment aesthetically pleasing and appreciated seeing the familiar CSU VTH represented in the application. Many commented on the cuteness of the pets, and when asked whether they would prefer cartoon-like or more realistic avatars, five out of six participants expressed their liking for the current cartoon-like avatars. They also appreciated the detailed background elements, with Participant 4 noting "I like the, and it's silly, but I think in one of the learning consults they have little Milk Bones in the back and I'm like, that's cute." The only negative feedback about aesthetics came from the rare instances where the environment glitched and turned completely black, as previously mentioned in Section 4.1.8.

4.2.3 Comparison to Traditional Methods

Participants unanimously felt that the virtual consultations represented an improvement over previous homework assignments on Canvas. They appreciated the ability to engage in real-time conversations with a client and receive immediate feedback on their grammar, including expla-

nations for any mistakes. This was a significant enhancement compared to the prior method of recording solo videos where they pretended to speak to a client. The main drawback of the video approach was the lack of interactive dialogue and the delayed instructor feedback, which could take weeks, often resulting in participants forgetting their initial responses.

Participants expressed enthusiasm about the potential to continue practicing their veterinary Spanish using the application in the future. However, they emphasized that certain technical issues such as speech recognition, grammar sensitivity, and the duration of consultations must be addressed before they would be willing to use it outside of class. Additionally, they expressed a desire for more diverse content to keep the scenarios fresh and engaging. Overall, while participants were excited about the application's prospects, they made it clear that resolving these concerns would be essential for them to fully utilize the tool.

4.2.4 Practical Application to Veterinary Training

When asked if the knowledge provided by the application applied to real-world scenarios, the participants said that it did and that the application provided a great opportunity to practice consultations without the pressure of talking to a real person. Participant 4 said, "It reminds you to always address the client, [...] this is preparing you for that."

However, participants said that it would be more applicable to a real-life consultation if the AI was not constantly correcting you on every spelling mistake made, or looking for specific answers because that is not how a human will act in a real-world consultation. or trying to transcribe ums and stuttering as words. They said that the app crashing and taking a long time makes it less applicable to a real-life consultations.

Chapter 5

Discussion

The findings of this study indicate that virtual learning environments can significantly enhance both language acquisition and practical diagnostic skills in veterinary education. Thematic analysis revealed key areas for improvement, including enhancing the usability of certain features, introducing new elements to boost engagement, and an overall positive reception of the interactive virtual experience.

5.1 Bugs to Be Resolved

5.1.1 Fix Bugs Related to Scrolling

The most critical improvement to navigation is fixing the bugs that cause the webpage to scroll or zoom while users navigate the environment. Possible solutions include disabling arrow key scrolling or ensuring the page fits the window, eliminating the need for scrolling or zooming.

5.1.2 Mitigate Loading Bugs and Crashes

The application currently takes a lot of processing power that participants stated could not be handled by their devices. This led to the program crashing frequently, long loading times, and lagging graphics & conversations. Some of these problems might be mitigated by employing graphic and performance optimization, such as adjustable graphics quality settings, or lowered frame rates depending on machine and network capabilities.

5.2 Improvements to Existing Features

5.2.1 Remove Strict AI-Enforced Grammar Rules on Conversation

The two most common complaints all participants made were that speech-to-text was inaccurate and AI was too picky about spelling and grammar. These two issues exacerbated each other,

with inaccurate speech-to-text leading to frequent complaints from the AI agent. The students wanted a more realistic conversational experience, and these issues led to clunky conversations in which participants repeated sentences constantly, got frustrated, lost track of the conversation, and gave up on speaking. While it's clear that speech-to-text algorithms have a lot of room for improvement, the participants expressed more interest in having more grace for making mistakes.

Since most of the participants shared the sentiment that they wished the conversation would continue after a spelling or grammatical mistake was made, but a few students also expressed gratitude for being corrected on their grammar, my proposed solution is to display a corrected version of their submitted sentence below the original sentence, with the changes highlighted as one participant recommended, but let the virtual client respond to what the participant submitted as someone would in a real-world conversation. Then, if the mistake is minor, the agent can infer the intended meaning, whereas if the mistake breaks comprehension, it can ask for clarification. This would not only make the interaction less frustrating for users but make it more similar to a real-life interaction that they are practicing for, with opportunities to try to explain their intended meaning as you would in real life if someone doesn't understand what you were trying to say. With this being the case, the quality of the speech-to-text is less important and more realistic because people often mishear and misunderstand people when they are talking in everyday conversation anyway. This solution would make the AI's corrections less intrusive and the inaccuracies of the speech-to-text less problematic.

5.2.2 Improve AI Responses

While there is no straightforward way to solve the issue of the AI contradicting itself or making nonsensical statements, there are strategies that could improve certain aspects of the conversation. For example, the AI could be instructed to give varying greetings at the start of each conversation, instead of repeating the same scripted introduction. Additionally, the AI could also be open to its information being incorrect. In the case of the nasal secretions example from Participant 2, the AI could say "I thought it looked clear" rather than "No, it's not green, Lulu's nasal secretion is clear!"

Many of these things could be adjusted in the parameters given to the AI, though it is impossible to foresee and prevent every awkward interaction.

5.2.3 Submit Text Easily

As the primary control for interacting with the AI agent, submitting responses should be a quick and intuitive process. Allowing users to submit their text by pressing the enter key is a convenient feature, however, participants occasionally encountered the unexpected scenario where pressing enter would start the microphone, which led to their response getting cleared instead of submitted. To improve the user experience, the enter key should be designated solely for submitting text to the AI agent and not for initiating voice recordings, which currently overwrites previously entered text. Assigning another key, such as the spacebar, to start the microphone recording could prevent this confusion. Notably, no participants intentionally pressed the enter key to start a voice recording.

5.2.4 Increase the Text Box Size

Participants frequently found the text box too small, particularly when trying to input full sentences or longer responses. The limited space made it difficult to see and review what they had written, leading to frustration and mistakes in their input. Many participants expressed a desire for a larger text box or one that could dynamically expand as they typed, allowing them to see their entire response at once. This would make it easier for users to edit their responses before submitting them.

5.2.5 Improve the Form

The consensus among participants was that the form was too wordy and repetitive. Many participants expressed a preference for bullet points, or short questions to guide users through the topics. One participant mentioned that they wished the form resembled those used in real veterinary clinics. While one participant stated they would prefer a step-by-step guide through the conversation, that approach does not align well with the unpredictable nature of a real conversation.

5.2.6 Clearly Indicate Saved Progress When Resuming

The current session-resuming feature leaves many participants uncertain whether their progress from the previous session is being retrieved, due to the lack of clear ques that indicate this. Without clear confirmation, users are left confused and unsure if they are continuing from where they left off or starting anew. To address this, it is recommended that when resuming a session, users should be returned to their previous position with the past conversation visible in the chat window.

5.3 New Features

5.3.1 Track the User's Progress Through the Application

Seven of the eight participants expressed a desire for some form of progress tracking through the application, with the most popular suggestion being the addition of checkmarks next to completed modules. The participants highlighted that progress tracking would not only help them stay organized, but also provide a sense of accomplishment, making it easier to gauge their learning journey. Other features proposed included a list of mastered topics or pets treated, progress bars visible during activities, and highlighting the activity "up next." This aligns with features in many other educational applications, where progress bars, checklists, and milestone indicators have been shown to boost motivation and engagement.

5.3.2 Add Easy Access to Language-Specific Characters

Many of the participants struggled with typing characters that are not used in English. Since keyboards are formatted for specific languages, this can pose a challenge when typing in a foreign language. Given that MeTabi is designed for practicing foreign languages, it's important to recognize that users may have difficulty typing language-specific characters. A simple solution, as suggested by Participant 6, would be to add on-screen buttons that users can click to insert the characters into the text box. This provides a convenient alternative for users who wish not to reconfigure their system keyboard settings. Additionally, Participant 2 expressed interest in a

built-in spell-check feature, which could assist users by offering suggestions for correct spellings in a dropdown menu.

5.3.3 Add Instructions and a Tutorial

While it was repeatedly expressed that the navigation was challenging for participants, most students didn't have many suggestions for improving the controls. Notably, three participants who struggled with navigation attributed it to a lack of experience playing video games, indicating a disadvantage for those unfamiliar with navigating virtual environments. Conforming navigational controls to common virtual environment navigation standards, such as what would be used in three-dimensional (3D) video games could be a way to make the experience easier for users who have prior experience in such environments. Many participants didn't fully understand all the available navigational controls in the application, with a couple of participants suggesting that adding instructions or a tutorial would be helpful.

5.4 Experience

5.5 Implications for Veterinary and Language Education

Participants said that they were more engaged and immersed in the MeTabi application than they had been in previous forms of homework and expressed a positive reaction to the aesthetics of the environment. Participants said that they preferred MeTabi over previous methods of language learning and could be interested in continuing to use the application in the future, as long as the bugs were fixed. Participants expressed that the virtual consultations in MeTabi felt relevant to real-world applications. As veterinary professionals are increasingly expected to interact with diverse populations, the ability to communicate effectively in another language is crucial. The integration of a virtual environment into veterinary education, particularly for language learning, offers a promising approach to bridging the gap between academic knowledge and practical application.

Chapter 6

Limitations

Despite these findings, there are limitations to consider. The sample size was relatively small, and all participants were from the same course in the same academic institution, which may limit the generalizability of the results.

Additionally, while the virtual environment simulated realistic veterinary interactions, future research should explore how students apply these skills in real-world clinical settings with Spanish-speaking clients. The comfort of virtual client interactions may not fully capture the complexity of real-world scenarios. It can often be much easier to talk to a computer than native human speakers. So how students handle communication challenges in live clinical environments, especially with actual Spanish-speaking clients, should be further explored.

Finally, the study did not explore in depth how different learning styles or individual preferences might influence user experience with the virtual environment. Some students might benefit more from highly structured interactions, while others may prefer more open-ended, exploratory scenarios. Understanding how the virtual tool can be adapted to different learning needs would further enhance its educational value.

Chapter 7

Conclusion

This study explores the potential of virtual learning environments to enhance language learning and professional skill development in specialized fields like veterinary medicine. By immersing students in a task-based context, virtual simulations can provide an innovative and effective way to prepare veterinary students for the challenges of working in multilingual and multicultural environments. This virtual environment was found to enhance participants' immersion and perceived engagement, while also providing practical knowledge for veterinary students learning Spanish. Participants viewed it as an improvement over previous homework and practice methods.

However, participant consensus was that improving the usability of certain features, most importantly improving the flow of conversation, explaining the navigation controls clearly, and implementing a way to easily track one's progress through the application, will be critical to reducing user frustration and optimizing the learning experience. Addressing the changes proposed here will enhance user engagement and make the application more accessible and effective for a wider range of users. Future research should continue to explore the integration of language learning technologies in professional education, especially in fields that require specialized communication skills.

Chapter 8

Future Work

While this research sheds light on the current state of technical language learning with virtual environments, there are plenty of future directions to explore. One possibility is to investigate how constant AI feedback on grammar and spelling impacts learning outcomes. This could involve asking students to complete two versions of a conversation: one where the AI provides continuous corrections and another without feedback. Researchers could then analyze the differences between the two experiences.

Another avenue for future research is to compare interactions with a virtual client like MeTabi against an in-person standardized client. By providing both clients with the same script or prompt, researchers could explore how conversation dynamics differ between virtual and real-life interactions.

Additionally, examining user preferences for avatar styles could be insightful. A study could compare participant interactions with the current cartoon-like avatars against more realistic-looking avatars to understand how aesthetic differences influence user engagement and immersion. The analysis could provide valuable feedback on whether the visual style impacts how seriously students take the interaction or how immersed they feel during consultations.

Further exploration could also focus on different modalities, such as AR and VR, to assess their potential for enhancing language learning experiences. By comparing user experiences, researchers could identify which modalities best support language acquisition and professional skill development.

Lastly, quantitative studies on long-term retention of Spanish and learning gains would be valuable. These studies could track participants over several months or even a year to measure how well they retain the technical language skills acquired through the virtual environment. By testing language proficiency at different intervals post-interaction, such research could shed light on the lasting impact of virtual learning environments compared to traditional methods.

Bibliography

[1] Spanish Speaking Countries | Overview & History.

[2] US Census Bureau. Nearly 68 Million People Spoke a Language Other Than English at Home in 2019. Section: Government.

[3] Ozzie F. Díaz-Duque. Communication barriers in medical settings: Hispanics in the United States. 1989(79):93–102, January 1989. Publisher: De Gruyter Mouton Section: International Journal of the Sociology of Language.

[4] Leandra J. Nogueira Borden, Cindy L. Adams, and Lynda D. Ladner. The Use of Standardized Clients in Research in the Veterinary Clinical Setting. *Journal of Veterinary Medical Education*, 35(3):420–430, September 2008.

[5] Juan Cendan and Benjamin Lok. The use of virtual patients in medical school curricula. *Advances in Physiology Education*, 36(1):48–53, March 2012. Publisher: American Physiological Society.

[6] Rafael Villena-Taranilla, Sergio Tirado-Olivares, Ramón Cózar-Gutiérrez, and José Antonio González-Calero. Effects of virtual reality on learning outcomes in K-6 education: A meta-analysis. *Educational Research Review*, 35:100434, February 2022.

[7] Ethan Seefried, Mariah Bradford, Swagatalaxmi Aich, Caspian Siebert, Nikhil Krishnaswamy, and Nathaniel Blanchard. Learning Foreign Language Vocabulary Through Task-Based Virtual Reality Immersion. In Constantine Stephanidis, Margherita Antona, Stavroula Ntoa, and Gavriel Salvendy, editors, *HCI International 2024 Posters*, pages 203–213, Cham, 2024. Springer Nature Switzerland.

[8] Thomas D. Parsons and Albert A. Rizzo. Affective outcomes of virtual reality exposure therapy for anxiety and specific phobias: A meta-analysis. *Journal of Behavior Therapy and Experimental Psychiatry*, 39(3):250–261, September 2008.

[9] Albert Skip Rizzo, Arno Hartholt, and Sharon Mozgai. Settling the Score: Virtual Reality as a Tool to Enhance Trauma-Focused Therapy for PTSD. In Grant J. Rich, V. K. Kumar, and Frank H. Farley, editors, *Handbook of Media Psychology: The Science and The Practice*, pages 187–213. Springer Nature Switzerland, Cham, 2024.

[10] Maria T. Schultheis and Albert A. Rizzo. The application of virtual reality technology in rehabilitation. *Rehabilitation Psychology*, 46(3):296–311, 2001. Place: US Publisher: Educational Publishing Foundation.

[11] A.A. Rizzo, J.G. Buckwalter, T. Bowerly, C. Van Der Zaag, L. Humphrey, U. Neumann, C. Chua, C. Kyriakakis, A. Van Rooyen, and D. Sisemore. The Virtual Classroom: A Virtual Reality Environment for the Assessment and Rehabilitation of Attention Deficits. *CyberPsychology & Behavior*, 3(3):483–499, June 2000. Publisher: Mary Ann Liebert, Inc., publishers.

[12] Christopher You, Tabitha Peck, Jacob Stuart, Alexandre Gomes de Siqueira, and Benjamin Lok. What my bias meant for my embodiment: an investigation on virtual embodiment in desktop-based virtual reality. *Frontiers in Virtual Reality*, 5, March 2024. Publisher: Frontiers.

[13] Gregory Guest, Kathleen M MacQueen, and Emily E Namey. Introduction to applied thematic analysis. *Applied thematic analysis*, 3(20):1–21, 2012.

[14] M. Kozlova. The benefit of immersive language-learning experiences and how to create them. *Retrieved from Canbridge: https://www. cambridgeenglish. org/blog/the-benefit-of-immersivelanguage-learning-experiences-and-how-to-create-them*, 2021.

[15] Celeste Kinginger. Enhancing Language Learning in Study Abroad. *Annual Review of Applied Linguistics*, 31:58–73, March 2011.

[16] Gustav Bøg Petersen, Giorgos Petkakis, and Guido Makransky. A study of how immersion and interactivity drive VR learning. *Computers & Education*, 179:104429, April 2022.

[17] Baron L. Savage and Haning Z. Hughes. How does Short-term Foreign Language Immersion Stimulate Language Learning? *Frontiers: The Interdisciplinary Journal of Study Abroad*, 24(1):103–120, August 2014. Number: 1.

[18] Dongsong Zhang, Lina Zhou, Robert Briggs, and Jay Nunamaker. Instructional video in e-learning: Assessing the impact of interactive video on learning effectiveness. *Information & Management*, 43:15–27, January 2006.

[19] IELTS IELTS and OET OET. The benefit of immersive language-learning experiences and how to create them.

[20] Kagan Buyukkarci. A critical analysis of task-based learning. *Kastamonu Journal of Education*, 17:313–320, January 2009.

[21] Jane Willis. *A Framework for Task-based Learning*. Intrinsic Books Ltd, October 2021. Google-Books-ID: 8MtGEAAAQBAJ.

[22] Sarah Garcia, Ronald Kauer, Denis Laesker, Jason Nguyen, and Marvin Andujar. A Virtual Reality Experience for Learning Languages. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*, CHI EA '19, pages 1–4, New York, NY, USA, May 2019. Association for Computing Machinery.

[23] Shiva Pedram, Sarah Howard, Ken Kencevski, and Pascal Perez. Investigating the Relationship Between Students' Preferred Learning Style on Their Learning Experience in Virtual Reality (VR) Learning Environment. In Tareq Ahram, Redha Taiar, Serge Colson, and Arnaud Choplin, editors, *Human Interaction and Emerging Technologies*, Advances in Intelligent Systems and Computing, pages 275–281, Cham, 2020. Springer International Publishing.

[24] Yu-Ju Lan. Chapter One - Immersion into virtual reality for language learning. In Kara D. Federmeier and Hsu-Wen Huang, editors, *Psychology of Learning and Motivation*, volume 72 of *Adult and Second Language Learning*, pages 1–26. Academic Press, January 2020.

[25] Alyssia M. Miller De Rutté. Virtual Reality, Medical Spanish, and the L2 Motivational Self System: A Case Study. *Technology, Knowledge and Learning*, 29(2):827–841, June 2024.

[26] George A. Miller and Patricia M. Gildea. How Children Learn Words. *Scientific American*, 257(3):94–99, 1987. Publisher: Scientific American, a division of Nature America, Inc.

[27] João Barreira, Maximino Bessa, Luciana C. Pereira, Telmo Adão, Emanuel Peres, and Luís Magalhães. MOW: Augmented Reality game to learn words in different languages: Case study: Learning English names of animals in elementary school. In *7th Iberian Conference on Information Systems and Technologies (CISTI 2012)*, pages 1–6, June 2012. ISSN: 2166-0735.

[28] Antigoni Parmaxi and Alan A. Demetriou. Augmented reality in language learning: A state-of-the-art review of 2014–2019. *Journal of Computer Assisted Learning*, 36(6):861–875, 2020. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/jcal.12486.

[29] W. G. Harless, G. G. Drennon, J. J. Marxer, J. A. Root, and G. E. Miller. CASE: a Computer-Aided Simulation of the Clinical Encounter. *Academic Medicine*, 46(5):443, May 1971.

[30] Pedro Guillermo Feijóo-García, Chase Wrenn, Alexandre Gomes de Siqueira, Rashi Ghosh, Jacob Stuart, Heng Yao, and Benjamin Lok. Exploring the Effects of User-Agent and User-Designer Similarity in Virtual Human Design to Promote Mental Health Intentions for College Students, May 2024. arXiv:2405.07418 [cs].

[31] Heng Yao, Alexandre Gomes de Siqueira, Megan L. Rogers, Sarah Bloch-Elkouby, Olivia Lawrence, Giuseppe Sarli, Adriana Foster, Serge A. Mitelman, Igor Galynker, and Benjamin Lok. The impact of scaffolded and non-scaffolded suicidal virtual human interaction training on clinician emotional self-awareness, empathic communication, and clinical efficacy. *BMC Medical Education*, 24(1):413, April 2024.

[32] Gary M. Olson, Susan A. Duffy, and Robert L. Mack. Thinking-Out-Loud as a Method for Studying Real-Time Comprehension Processes. In *New Methods in Reading Comprehension Research*. Routledge, 1984. Num Pages: 34.

[33] Heather L. O'Brien, Paul Cairns, and Mark Hall. A practical approach to measuring user engagement with the refined user engagement scale (UES) and new UES short form. *International Journal of Human-Computer Studies*, 112:28–39, 2018.