

THESIS

LEXVET-ESP: DEVELOPING A NEW VOCABULARY TEST AND THE INTEGRATION
OF VR FOR SPECIALIZED VOCABULARY ACQUISITION

Submitted by

Paula Izquierdo García

Department of Languages, Literatures and Cultures

In partial fulfillment of the requirements

For the Degree of Master of Arts

Colorado State University

Fort Collins, Colorado

Summer 2025

Master's Committee:

Advisor: Alyssia Miller De Rutté

Shannon Zeller
Lily Edwards-Callaway

Copyright by Paula Izquierdo García 2025

All Rights Reserved

ABSTRACT

LEXVET-ESP: DEVELOPING A NEW VOCABULARY TEST AND THE INTEGRATION OF VR FOR SPECIALIZED VOCABULARY ACQUISITION

The number of Spanish speakers in the U.S. continues to increase, which leads to a growing population of individuals who do not speak English as their primary language, making access to various services, including veterinary care, difficult. To address this issue, Languages for Specific Purposes (LSP) courses have been developed to train (future) professionals to speak in their client's preferred language. Alongside the rising popularity of LSP courses, technological advancements have also been implemented in language education. Virtual reality (VR) platforms and artificial intelligence (AI) are two examples of relevant tools in second language learning as they allow for the integration of immersive and interactive experiences. This thesis aimed to combine LSP, particularly in Spanish for Veterinary Medicine, and technology to understand the potential effects on language learners' vocabulary acquisition and retention. The thesis was divided into two projects. The first project focused on developing an assessment tool to test learners' receptive vocabulary knowledge, defined as words that learners can recognize even if they cannot yet define or use them in different contexts. This assessment builds upon previous studies, which created validated and reliable instruments for measuring vocabulary and proficiency levels. A specialized vocabulary test for Veterinarian Spanish (known as the LexVet-Esp) was developed and validated. The second project implemented a VR/AI platform as part of a Spanish for Veterinary Medicine course and used the LexVet-Esp to assess the impact of the technology on vocabulary acquisition and retention and whether explicit or implicit vocabulary

exercises contributed to language development. To examine this, students completed the vocabulary test four times. The first test, administered during week nine of a sixteen-week semester, served as the pre-test as participants had not yet interacted with the VR/AI platform. Students took the test again in week twelve after completing three weeks of explicit vocabulary exercises on the platform and again in week fifteen after completing implicit vocabulary activities. Finally, a fourth test was conducted three months after the end of the course to determine long-term vocabulary retention. The results did not show significant differences across the tests, which may have been influenced by the short length of the study (six weeks) and the small sample size (n=15). Despite the lack of statistically significant differences across the tests, results indicated varying complexities related to the vocabulary acquisition and retention processes. Pedagogical implications and future research opportunities are discussed.

TABLE OF CONTENTS

ABSTRACT.....	ii
CHAPTER 1: INTRODUCTION.....	1
1.1 Chapter Overview.....	4
CHAPTER 2: LITERATURE REVIEW.....	6
2.1 Language for Specific Purposes.....	6
2.1.1 Spanish for Specific Purposes.....	8
2.1.1.1 Spanish for Doctor of Veterinary Medicine Students.....	11
2.1.2 Language Needs Analysis.....	13
2.1.3 Task-Based Language Teaching.....	15
2.1.3.1 LNAs, TBLT, and LSP.....	17
2.2 Technology in Language Education.....	18
2.2.1 VR and Language Education.....	20
2.2.1.1 Grounded Cognition and Vocabulary Acquisition.....	22
2.2.2 AI and Language Education.....	23
2.2.3 VR and AI in LSP.....	24
2.2.4 VR and AI in the Veterinary Field.....	25
2.3 Vocabulary Acquisition in Language Learning.....	26
2.3.1 Vocabulary Assessments.....	28
2.4 Research Questions.....	31
CHAPTER 3: THE DESIGN AND VALIDATION OF THE LEXVET-ESP TEST.....	32
3.1 Introduction to Project 1.....	32
3.2 Methodology.....	32
3.2.1 Materials.....	32
3.2.2 Procedure and Participants.....	35
3.2.3 Data Analysis.....	36
3.3 Results.....	37
3.3.1 Point-biserial Correlation.....	38
3.3.2 Item Response Theory.....	39
3.3.3 Comparisons across proficiency levels.....	42
3.4 Discussion.....	43
3.4.1 Comparison of the LexVet-Esp with LexITA and Lextale-Esp.....	43

3.4.2	Addressing the Limitations of Point-Biserial Correlations with IRT	44
3.4.3	Differentiation of Proficiency Levels	45
3.4.4	Limitations	46
CHAPTER 4: MEASSURING VOCABULARY ACQUISTION WITH LEXVET-ESP IN A VIRTUAL REALITY ENVIRONMENT.....		47
4.1	Introduction to Project	47
4.2	Methodology	48
4.2.1	Participants and Context	48
4.2.2	Materials	48
4.2.3	Procedure	55
4.2.4	Data Analysis	56
4.3	Results.....	57
4.3.1	Assessing Performance Consistency and Retention Over Time.....	62
4.4	Discussion.....	63
4.4.1	Impact of VR and AI on Vocabulary Acquisition	63
4.4.2	Implication for Educational Practice	66
4.4.3	Limitations and Future Research	67
CHAPTER 5: DISCUSSION AND CONCLUSION		68
5.1	Validation of LexVet-Esp.....	69
5.2	Considerations for VR, AI, and Vocabulary Retention	69
5.3	Future Research	70
REFERENCES		72
APPENDICES		88
Appendix A: Complete list of 180 terms used to develop the LexVet-Esp.....		88
Appendix B: Part 1, Instructions, and Example Included in the LexVet-Esp		90

CHAPTER 1: INTRODUCTION

The world is becoming increasingly globalized, leading to a rise in international migration. This phenomenon often results in people relocating to countries where they do not speak the primary language. This process happens in the United States (U.S.), and U.S. immigrants often find employment in a variety of sectors, where language barriers can significantly impact their daily lives and work experiences. An examination of the U.S. occupational data shows that 36.2% of immigrants are classified as being in management, business, sciences, and arts roles, while 21.3% are in the service sector (US Census Bureau, 2024c). When considering the educational level of this population, it is important to note that only 14.9% have a graduate or professional degree, which is in contrast to the 25.6% who do not hold a high school diploma or equivalent (US Census Bureau, 2024c).

Language barriers can be especially pronounced in service-oriented professions, where clear communication is essential for successful interactions. One critical area where communication gaps manifest is in healthcare settings, particularly in patient-doctor interactions. Language differences can lead to inaccurate patient assessments, misunderstandings about treatment plans, and ultimately, poorer health outcomes. Similarly, language-related challenges extend to other professions that require direct engagement with clients, including veterinary medicine. In recognition of these challenges, there is a growing emphasis on addressing language barriers across professional fields with increasing focus in language education on teaching Languages for Specific Purposes (LSP). This field of study aims to equip professionals with the language and intercultural skills necessary for their workplace, which will ensure effective communication with diverse client population. In the context of this study, the focus is on

teaching Spanish for Veterinary Medicine in which the goal is helping veterinary students in the U.S. establish better rapport with their Spanish-speaking clients through the Spanish language, as the Hispanic population in the U.S. continues to rise.

To better understand this study's importance, it is essential to present current statistics surrounding immigration within the U.S. According to data provided by the US Census Bureau (2024b), the number of foreign-born people has increased to over 45.3 million, which equals 13.7% of the total U.S. population. A significant proportion of this demographic is concentrated in California, New Jersey, New York, and Florida. According to the U.S. Census (2024a), approximately 65.2 million people in the U.S. (19.5% of the population) identify as Hispanic. As shown in the American Community Survey five-year estimates, the U.S. population is projected to be 27% Hispanic/Latino by 2060 (US Census Bureau, 2023). Spanish is the second most spoken language in the U.S., with 22.5% of the population communicating in a language other than English and 13.7% using Spanish as their primary language (U.S. Census Bureau, 2024a). As Spanish continues to grow in usage, the need for Spanish-proficient professionals across industries, including veterinary medicine, is becoming increasingly apparent.

As the Hispanic population has grown, so has the rise in pet ownership. Veterinary professionals frequently engage with Hispanic clients, many of whom may prefer to communicate in Spanish when discussing the health and care of their animals. As reported by Larkin (2024), there has been a significant increase in the number of pets within households when comparing data from 1996 to 2024. The growing number of pet owners reinforces the need for accessible veterinary care that accommodates linguistic diversity. According to Brown (2023), 62% of the U.S. adult population owns a pet, and 66% of Hispanic individuals report

owning at least one pet. The canine population alone in 2024 was 89.7 million in the U.S., underscoring the broad presence of pets in American households.

Even though there has been a high level of pet ownership, there has been a drop in cost associated with veterinary care for dogs. For example, the average cost for routine checkups or preventive care, which represents 80% of visits in 2024, was \$147 per visit compared to \$190 in previous years (Larkin, 2024). This trend, combined with the growing Hispanic population, highlights the increasing accessibility of veterinary care. However, as more Spanish-speaking pet owners seek routine services, the need for improved communication between veterinary professionals and their clients becomes even more essential. Despite this growing demand, only 10% of veterinary professionals report proficiency in Spanish, although it is unknown how proficiency was assessed to arrive at this statistic (Hopkins, 2023). A range of solutions to increase the number of Spanish-speaking veterinary professionals has been proposed and include the development of specialized courses or certificate programs designed to teach Spanish to veterinary students across the U.S. (Hopkins, 2023). These initiatives aim to equip veterinary students with the language skills necessary to communicate effectively with Spanish-speaking clients, ultimately improving the quality of care provided.

In addition to specialized courses and certificate programs, advancements in technology, including virtual reality (VR) and artificial intelligence (AI), are being used to provide language learners immersive and interactive learning experiences in real-world environments. VR allows learners to practice a foreign language in authentic, real-world scenarios, which enhances engagement and motivation (Chen et al., 2022; Özgün & Sadik, 2023). Similarly, AI can facilitate training in the language and give students real-time feedback on writing and pronunciation (Macinska & Vinkler, 2021; Wetherbee, 2023). These technological innovations

have the potential to help learners improve their linguistic skills (Sun, 2023), which has the potential to revolutionize language learning, making it more accessible and effective for learners. Therefore, the purpose of this thesis was to integrate a VR and AI platform into a Spanish for Veterinary Medicine course to investigate the effects on second language acquisition with a focus on vocabulary learning and retention.

1.1 Chapter Overview

The remaining chapters in this thesis present the theoretical framework, methods, and practical applications of the study. Chapter 2 provides a comprehensive review of previous literature relevant to the study. It explores LSP, with a focus on Spanish for veterinary professionals, alongside methodologies such as the Language Needs Analysis and Task-Based Language Teaching. Additionally, the chapter examines the role of technology in language education, discussing VR, AI, and their applications within the LSP and veterinary fields. The chapter continues with a review of vocabulary retention in language learning and assessment methods used to measure vocabulary knowledge. It then concludes with the research questions that guide the thesis.

Chapter 3 presents Project 1, which focuses on the creation and validation of a vocabulary assessment tool for Spanish for Veterinary Medicine, known as the LexVet-Esp. The methodology section outlines materials, procedures, and data analysis techniques, including point-biserial correlations and Item Response Theory to ensure the test's reliability and validity. The results present the discriminative power of test items, comparisons across proficiency levels, and how the LexVet-Esp aligns with existing assessments in other languages and areas.

Chapter 4 details Project 2, which examines vocabulary retention in a VR/AI-enhanced learning environment using the LexVet-Esp as an assessment tool. The methodology describes participant recruitment, study context, materials, and procedures for integrating the VR/AI technology into a Spanish for Veterinary Medicine course. The results assess student performance and retention across multiple test points and the effects of explicit vs. implicit vocabulary instruction.

Chapter 5 synthesizes findings from both projects, offering a more extensive discussion on the thesis as a whole. It revisits LexVet-Esp's validation and discriminative power, explores trends in VR-assisted vocabulary retention, and highlights methodological limitations. The chapter concludes by summarizing the study's contributions and suggesting areas for future exploration, particularly regarding long-term vocabulary retention and the integration of learning technologies in LSP instruction.

CHAPTER 2: LITERATURE REVIEW

This chapter provides an analysis of the most relevant literature and theoretical frameworks that are related to this study. It is organized into four main sections, starting with an overview of LSP before focusing on the subfield of Spanish for Specific Purposes (SSP). Then, a summary of two connected areas is presented. The first is an explanation of a Language Needs Analysis (LNA), which is an LSP research methodology, and the second is an overview of Task-Based Language Teaching (TBLT), a language teaching methodology associated with the results of an LNA. The second main section centers on the implementation of VR and AI in language education with emphasis on the use of these technologies in LSP. The third main section is focused on the area of vocabulary acquisition and retention in the field of Second Language Acquisition (SLA). The fourth and final section presents the research questions associated with this project.

2.1 Language for Specific Purposes

To better understand what LSP is, specialty languages have to be defined. Gómez de Enterría (2009) defines specialty languages as the languages used in the fields of science, technology, and the professions. These languages are used to transmit specialized knowledge, and Gómez de Enterría (2009) stated that all specialty languages can share common linguistic and functional characteristics. The field of LSP emerged as a way to address the language training of professionals who would use specialty language in their careers. Swales (2000) stated that the beginnings of LSP can be traced back to 1964 when a relationship between linguistic analysis and educational materials was first established. Initially, there was little emphasis on language educators needing expert knowledge of the specialized fields, and instead, LSP

practitioners were skilled at conducting basic descriptions of target discourses. As the LSP field grew, there was a need for a more refined approach to research and teaching LSP, and the focus shifted from basic language skills to discourse and genre analysis during the 1980s. Nowadays, LSP maintains a unique relationship with other branches of applied linguistics with close connections to discourse analysis and pragmatics with connections to business and technical communication, translator training, language assessment, and communicative language teaching (Swales, 2000). LSP's status as a discipline and profession varies globally with limited presence in U.S. graduate programs due to its separation from language acquisition, language teaching methodology, psycholinguistics, and sociolinguistics (Swales, 2000).

In the context of the U.S., LSP research faces several challenges, including the need for interdisciplinary integration, the difficulty of collecting authentic and relevant data, and the development of effective assessment tools for specific professional skills. Collaboration between language departments and field-specific programs has become important as “land grant institutions...working in cohort with language departments are particularly well suited to spearhead domain-specific language programs that are methodologically sound and are based on established best practices within the LSP field” (Zeller & Velázquez-Castillo, 2018, p. 295). Adapting to new technologies and incorporating them into LSP curricula presents ongoing challenges. Other challenges to be considered are the global variability of its status (what is the status/recognition in different countries) and the recognition of LSP as a discipline (Lafford, 2012). Additionally, there is often a lack of trained LSP practitioners and researchers with expertise in SLA.

Sánchez-López et al. (2017) conducted research in the U.S. in the context of higher education to determine the main areas of LSP that are relevant for research amongst LSP

scholars. Through interviews with researchers, Modern Language Association (MLA) members, and department chairs from various institutions in the U.S., Sánchez-López et al. (2017) found the areas of greatest interest for LSP research were business, culture, translation, academic purposes, service learning/community engagement, and health. These findings highlight the diverse professional domains in which language proficiency is essential, reinforcing the growing demand for specialized language instruction tailored to specific career fields.

Due to the increasing interest in applying language study to professional contexts, these types of courses appeal to many students (Pastor Cesteros, 2013). Students who enroll in LSP classes typically have a different motivation compared to general language learners as LSP courses have lexical, grammatical, and textual properties that are also different from general language instruction. To align with professional objectives, research suggests that LSP courses integrate meaningful material, authentic texts, and goal-oriented activities designed for real-world application (El Arbaoui, 2024; Nepravishtha & Roseni, 2014). In response to this need, educators have developed specialized courses targeting specific field, such as Chinese or Korean for Business, Spanish for Nurses, English for Home Care, Mandarin for Tourism, Legal Arabic, among others (Helms et al., 2023; Trace et al., 2015).

2.1.1 Spanish for Specific Purposes

There has been an increasing interest in learning Spanish for Specific Purposes with the most popular areas being science and technology, law, medicine, and tourism (Pastor Cesteros, 2013). The growth of the field of SSP is characterized by an emphasis on global context and collaboration. Significant contributions from Europe and Latin America highlight the importance of communication, connections, and collaboration among SSP scholars and practitioners (Lafford, 2022). Yu et al. (2020) stated that interdisciplinary work is indispensable in order to

have authentic and relevant LSP courses as through collaboration, LSP instructors can become more familiarized with the specialty area, and the content expert can become more aware of the importance that language and culture have in their discipline. The interdisciplinarity of this collaboration can occur at different points of the process, such as during course design or the delivery of the course, and it can be an intra-institutional or extra-institutional partnership (Yu et al., 2020).

SSP has become an integral part of Spanish curricula in universities, particularly in the U.S., where there is a focus on developing students' communicative competence to address the needs of marginalized Spanish-speaking communities (Lafford, 2024). The integration of SSP in language curricula aims to provide students with the practical language skills necessary for specific professional contexts, enhancing their employability and effectiveness in diverse fields. In the U.S., the fields that are most researched are related to Science, Technology, Engineering, and Math (STEM), healthcare, business, education, and agriculture (Helms et al., 2023; Lafford, 2024; Miller De Rutté et al., 2024; Pérez, 2021; Salazar et al., 2024; Zeller et al., 2016).

Institutions are developing specialized SSP curricula and conducting extensive research on curriculum design and effective teaching methodologies (Pérez, 2021). Zeller & Velázquez-Castillo (2018) provide a deeper contextualization of the need for the development of SSP programs. In their study, they offer a review of how various institutions have addressed SSP offerings, and they discuss how there exists a cohort of undergraduate students eager to participate in a non-traditional SSP area, which is Spanish for Animal Health and Care. The authors delineated the steps undertaken to establish an undergraduate certificate program in this area. They conducted an LNA, identified tasks (i.e., illness treatment and care, health histories, preventative care, among others), gathered additional information about these tasks (i.e.,

language functions and samples), and created a scaffolded sequence for the identified tasks to facilitate the development of a curriculum. Additionally, the study articulated the critical components that must be considered during the creation of SSP curricula. These components included identifying linguistic characteristics, grammatical forms and structures, lexical attributes, and sociocultural factors that influence professional communication (Zeller & Velázquez-Castillo, 2018).

Despite the advancements in SSP, challenges, such as the need for qualified instructors, integration of technology, and accurate assessment tools, persist (Czerkawski & Berti, 2020). SSP practitioners find themselves in a challenging position, caught between the demands of the language and the extensive diversity of scientific disciplines they are expected to address. They are tasked with teaching a specialized language, even though they often lack familiarity with the scientific fields and/or the linguistic tools required to transmit information related to those fields, which leaves them struggling to bridge the gap between two domains they might not fully command (Dodigovic, 1993). Because of the great diversity of disciplines, in most cases, each individual instructor has developed their own approach to teaching.

To address these challenges, there has been a call for stronger connections between LSP scholars and practitioners. Professional development in SSP is supported through conferences and publications with notable organizations, such as the Asociación Europea de Lenguas para Fines Específicos (AELFE) and the Congreso Internacional de Español para Fines Específicos (CIEFE), playing key roles in advancing the field (Lafford, 2022). These efforts highlight the dynamic nature of SSP and the ongoing commitment to its advancement through research, education, and professional collaboration.

2.1.1.1 Spanish for Doctor of Veterinary Medicine Students

“An awareness of Hispanic culture and language is becoming increasingly necessary within the US borders, especially for those who work directly with a Spanish-speaking workforce” (Zeller & Velázquez-Castillo, 2018, p. 290), and as previously mentioned, the area of Spanish for Animal Health and Care is growing at the undergraduate level. Similarly, the SSP field has begun to train future veterinary professionals at the graduate level in the Spanish language that is needed to communicate with Spanish-speaking clients. One initial attempt was carried out by Graves (2014, as cited in Zeller & Velázquez-Castillo, 2018) in which a bilingual workshop on cow reproduction was offered; however, this type of workshop did not focus on communication. Instead, training focused on providing translation of terms and phrases and did not focus on grammatical patterns or pragmatics of the language. Moreover, this training presented only one point of view, the English-speaking perspective, and did not consider speakers from other perspectives, such as the Spanish-speaking community. The lack of the Spanish speaker’s voice reinforces the privilege of dominant cultures while disadvantaging minority groups, and so, cultural competence must also be considered in these course offerings as it is only then that importance will be given to examining and addressing biases (Zeller et al., 2023).

In another study, Landau et al. (2015) sent a questionnaire to students at all veterinary schools in the U.S. to understand students’ perceptions of the use and need for Spanish in the veterinary field. Students were asked about their experience in and ability with Spanish as well as the topic of preparedness to communicate in medical terms. Results from this study indicated that there was a gap between students’ general and medical Spanish proficiency, which resulted in many of them feeling unprepared to give medical information to Spanish-speaking clients.

Moreover, some students, who self-identified as informal interpreters, did not consider themselves fluent in Spanish, which highlights a potential discrepancy between perceived and actual language skills. On the other hand, the need for Spanish in the veterinary field was not apparent for many of the students surveyed, and the ones who used Spanish in this setting did not stop to think about the challenges and difficulties that non-English speakers experience (Landau et al., 2015). As a result, Landau et al. (2015) stated that “there is room to improve professional communication competencies and diversity/multicultural awareness as identified by the North American Veterinary Medical Education Consortium” (p. 330).

Nowadays, there are U.S. institutions that provide Spanish language courses within their Doctor of Veterinary Medicine (DVM) curriculum. For example, Texas A&M University offers a course titled “Medical Spanish”, which is aimed primarily at teaching foundational Spanish skills to facilitate engagement with Spanish-speaking clientele (Zeller et al., 2023). Another institution is the University of California, Davis, which has integrated Spanish language courses into the DVM offerings in collaboration with the Spanish and Portuguese department, and Purdue University offers Spanish language instruction during lunch breaks to DVM students (Zeller et al., 2023). An additional initiative within the field of Animal Sciences, although not directly affiliated with DVM offerings, is the collaborative program established between Texas Tech University, North Carolina State University, and Tarleton State University (Salazar et al., 2024). This program consists of three courses in Spanish administered across three consecutive semesters. The sequence of courses starts at a basic level, concentrating on vocabulary and grammatical structures, and advances to more complex tasks at an intermediate level in the second course. The final course covers discussions and simulations that reflect real-world scenarios related to agriculture and animal care in Spanish.

Another study related to the Spanish for Veterinarians Language Program (SVLP) at a large western U.S. university (Forehand et al., 2023). Before developing the program, the researchers conducted a survey to determine if an eventual certificate in Spanish at the DVM level was of interest to the student body. A total of 791 students from across the U.S. responded to the survey, 275 of whom were enrolled at the researchers' institution. From this survey, it was found that there was great motivation for this type of coursework and that students were even willing to pay for classes out of pocket. Participants mentioned that their motivation to take these courses stemmed from the desire to help them improve client relations, increase confidence, and understand or potentially improve the animal's health. The SVLP was designed based on an extensive LNA and using TBLT, both of which will be described in the next section, to focus on specific communicative tasks in Spanish so that students could develop the relevant language that they would need to use in their future careers (Zeller et al., 2023).

2.1.2 Language Needs Analysis

Chambers (1980) originally explained that “needs analysis should be concerned with the establishment of communicative needs and their realizations, resulting from an analysis of the communication in the target situation” (as cited by Basturkmen, 2010, p. 18). Basturkmen (2010) revised Chambers' ideas 30 years later and stated that

needs analysis in [English for Specific Purposes; ESP] refers to a course development process. In this process the language and skills that the learners will use in their target professional or vocational workplace or in their study areas are identified and considered in relation to the present state of knowledge of the learners, their perceptions of their needs and the practical possibilities and constraints of the teaching context (p.19).

There are several advantages to doing this type of analysis, and one of those advantages is the fact that LNAs employ multiple sources of information, incorporating insights from students, professionals, and the target community who directly engages with the specialized language in practice (Malicka et al., 2019). One example of an LNA in the animal sciences field is work done by Zeller and Velázquez-Castillo (2018), who conducted a comprehensive needs analysis centered on the linguistic demands of the Spanish-speaking workforce in livestock environments. Their study involved observations of livestock work routines, interviews with practitioners and clients, and discussions with veterinarians and a farm manager, ensuring that the analysis reflected the real-world communication needs of professionals and the communities they serve. These analyses resulted in the development of a program for future professionals working with livestock farms and establishments. Additionally, the analysis of LNA data puts language tasks at the center. According to Candlin (1987), a task is “one of a set of differentiated, sequentiable, problem-posing activities involving learners and teachers” (as cited in Robinson, 2011, p. 6). Van den Branden (2006) adds that a task is “an activity in which a person engages to attain an objective, and which necessitates the use of language” (p. 4). Through the integration of perspectives from the target language community, a comprehensive and deep understanding arises regarding the specific types of language tasks learners must develop in order to meet the target language community members’ needs (Long, 2005). Therefore, it is through the LNA process that it will be possible to “relate instructional goals, processes, and practices to real-life performance outside the classroom” (Malicka et al., 2019, p. 79).

Another example of a comprehensive task-based LNA related to Spanish for Veterinarians at the DVM level involved observations at veterinary clinics in Colorado, Texas,

and Colombia, along with other data collection methods, including interviews with veterinary professionals and their Spanish-speaking clients (Zeller et al., forthcoming). This research aimed to improve access to veterinary care for Spanish-speaking pet owners by identifying the clients' communication needs and developing a curriculum tailored to these needs. The findings informed the creation of a Spanish language program that equips veterinarians with the necessary Spanish language skills to engage effectively with their clients. This program consists of five courses with a total of nine credits. Four of these courses are centered in the veterinary wellness appointment and includes taking the health history, relaying diagnostics, and discussing the treatment plan. All courses are conducted in Spanish. The fifth course, which is one credit, is focused on cultural awareness and access to care, and it is delivered in English. The analysis from the LNA also indicated the entrance proficiency level required for the program, which is a minimum of novice high on the American Council on the Teaching of Foreign Languages (ACTFL) proficiency scale. The proficiency level increases with each course so that veterinarians can build trust and develop stronger relationships with their clients, thereby improving the quality of care provided to the animals.

2.1.3 Task-Based Language Teaching

TBLT is the teaching methodology associated with the LNA research methodology. Ellis (2003) stated that “TBLT is an approach based on interactive and communicative tasks aiming at involving learners in meaningful communication and interaction enabling them to acquire linguistic structures as a result of engaging in authentic use” (as cited in Khatib & Dehghankar, 2018, p. 5). TBLT was developed out of the Communicative Language Teaching approach (Motlagh et al., 2014), which emphasizes interaction and communication as the primary goals of learning a new language (Qasserras, 2023). The core concept of the TBLT approach centers on

the task itself (Richards & Rodgers, 2001). El Arbaoui (2024) stated that the use of tasks encourages “effective and comprehensive language exposure and usage” (p. 253).

TBLT has been widely studied in SLA, and research has shown that tasks aid in several SLA processes, such as understanding input and producing output, as tasks provide students with opportunities to engage in meaningful interaction (Long, 1985; Robinson, 2011). Through task-based activities, learners are encouraged to produce language output, aligning with the Output Hypothesis (Swain, 1985), which suggests that producing language helps learners notice gaps in their knowledge and improve their language skills. Additionally, tasks facilitate a focus on form by drawing learners’ attention to specific linguistic forms while, at the same time, maintaining a focus on meaning (Ellis, 2003). This approach promotes negotiation of meaning, allowing learners to clarify misunderstandings and enhance their comprehension and production skills (Skehan, 1998).

TBLT considers numerous philosophical positions and empirical traditions from different fields, including education, applied linguistics, and psychology. Important considerations of this methodology relate to experiential learning, student-centeredness, and a process-oriented approach to syllabus creation (Nunan, 2014). These three characteristics mean that there is a focus on how learners acquire language skills and the strategies they use rather than just the final product or specific language items that students should master. It also proposes that when language is used to accomplish meaningful tasks, it significantly enhances the learning process. Furthermore, this approach highlights that language becomes more beneficial to learning when it holds personal significance or relevance for the learner (Motlagh et al., 2014). These principles underscore the importance of practical, purposeful language use in educational settings.

2.1.3.1 LNAs, TBLT, and LSP

LSP courses should be meticulously designed using evidence-based methodologies, such as the LNA, to address the language needs required by professionals in a certain domain, which also stresses the importance of adopting TBLT to replicate the real-world tasks that will be carried out outside of the classroom (Naudi, 2023). This approach allows learners to consistently engage with the language in both oral and written forms through activities focused on completing the task in its fullest, often leaving aside grammatical perfection (Cédric, 2021; Naudi, 2023). Investigations conducted by Hattani (2020), Georgy (2023), and Nazari (2020) outline both advantages and disadvantages associated with the incorporation of TBLT and LSP and are discussed next.

Among the advantages of combining TBLT in LSP courses are the facilitation of a significant contextual framework for learners as well as the promotion of a student-centered classroom environment. This approach enhances student engagement and contributes to a more stimulating and less repetitive classroom atmosphere (Hattani, 2020). Most of these instructional sessions are conducted in the target language so that students are immersed in the L2 during the class (Nazari, 2020). Due to the student-centered nature of the class, many activities are completed through collaborative engagement in pairs or groups to foster both cooperative and collaborative learning (Hattani, 2020). Research has shown that students find the TBLT nature of their LSP courses to be more engaging than general language courses, and they report an increased sense of self-efficacy when using their L2 (Hattani, 2020). Similarly, students articulate that this pedagogical approach allows them to identify more targeted learning objectives, providing them with better clarity regarding the skills and knowledge they are expected to develop during these classes (El Arbaoui, 2024). Finally, students expressed a desire

for instructors to adapt and implement this instructional methodology in other disciplines, including those not directly associated with language acquisition (El Arbaoui, 2024).

Some disadvantages of incorporating TBLT and LSP include the time investment required for lesson planning and finding authentic materials, which may not be widely available or adapted for use by language learners (Georgy, 2023). Additionally, instructors must be aware of their role as facilitators of information while simultaneously monitoring and providing feedback without disrupting the tasks or obstructing communication (Georgy, 2023; Hattani, 2020). However, instructors have noted an increase in students' motivation and self-esteem when employing this pedagogical approach, as it enables students to notice their relevance in real-world contexts (Hattani, 2020).

2.2 Technology in Language Education

In this next section, the role of technology in language education, at large, and in LSP are discussed. Technological instruments in the classroom should act as collaborators rather than obstacles to allow for the successful integration of supplementary resources and activities, and there are many possibilities to do so. As noted by Chacón Medina (2007), “the greatest potential of new information and communication technologies (ICTs) is derived from the capabilities of manipulation, storage, and distribution of information in an easy, fast, and accessible way for all people” (p. 25). ICTs can increase access to knowledge while optimizing educational processes. This characteristic makes technology indispensable in education, and the incorporation of technology into L2 teaching is becoming progressively common to help promote immersion in authentic contexts.

Technological advancements have brought a variety of tools and platforms that make language learning more engaging and effective (Mayer, 2014). Some of the most popular include multimedia tools, mobile apps, digital games, augmented reality (AR), and VR (Blyth, 2018; Godwin-Jones, 2011; Ibáñez & Delgado-Kloos, 2018; Pachler et al., 2010; Reinders & Wattana, 2015). Multimedia tools, such as videos and audio recordings, help learners improve their listening and speaking skills by exposing them to authentic language use (Pachler et al., 2010). Multimedia tools also help improve language comprehension and retention as they offer a variety of input types, such as audio, visual, and textual (Mayer, 2009). Apps like Duolingo and Babbel use smartphones and tablets to offer flexible, on-the-go learning opportunities, especially for learners with busy schedules or limited access to traditional classrooms (Kukulska-Hulme & Shield, 2008; Reinders & Hubbard, 2013). Mobile technologies use adaptive learning systems that personalize instruction to meet individual learners' needs while offering tailored feedback and practice, which boost learning efficiency. These apps also provide interactive exercises and gamified experiences that keep learners motivated (Godwin-Jones, 2011). Digital games can create immersive environments for practicing language through problem-solving and storytelling (Reinders & Wattana, 2015), which increases learner engagement and motivation and leads to higher participation rates and greater persistence in language study (Sykes & Reinhardt, 2012). Another example of this idea is found in AR as AR overlays digital information onto the real world to help learners engage with vocabulary and phrases in context (Ibáñez & Delgado-Kloos, 2018), while VR provides immersive virtual environments where learners can practice language skills (Blyth, 2018). AR and VR deliver contextualized practice, which helps students transfer language skills to authentic communicative environments (Holden & Sykes, 2011). Additionally, digital games and VR can support collaborative learning by allowing learners to interact and

communicate in virtual spaces to enhance language practice and foster social interaction (Thorne, 2008).

2.2.1 VR and Language Education

VR is one of the main technologies in this study, and this next section will focus on VR use in language education. VR is defined as “a simulation of a three-dimensional virtual environment, generated by a computer, in which the individual can engage with the aforementioned environment” (Peixoto et al., 2021, p. 48952). VR systems can be divided into three distinct categories: non-immersive, semi-immersive, and immersive. The principal differentiation among these categories is based on the degree of immersion experienced by the user, as well as the cost associated with these categories. As the level of immersion increases, there is a corresponding need for specialized equipment and advanced technological infrastructure to enable a more authentic and interactive environment (Peixoto et al., 2021), and all three levels of immersion have appeared in language teaching (Lin & Lan, 2015).

Immersive VR settings can replicate real-world situations, enabling students to develop language competencies within authentic contexts that are otherwise challenging to recreate in conventional classrooms. This type of VR can enhance speaking and listening skills as learners can practice conversations in a safe space, which helps them build both confidence and proficiency (Lin & Lan, 2015). In the world of immersive VR, new concepts have emerged, like the idea of the metaverse. According to Anaconda Ortiz et al. (2019), metaverses are characterized as “virtual worlds that allow users to let their imagination run free” (p. 62). Within educational frameworks, metaverses are described as “real-time immersive 3D simulated environments whose ecosystem is well-suited to incorporate audiovisual notifications, resulting in impressive configurations for formative or pedagogical spaces” (Barráez-Herrera, 2022, p. 16). The

transition towards immersive digital learning environments emphasizes the potential of VR technologies in augmenting student engagement and promoting authentic learning experiences.

The effectiveness of VR in the field of L2 learning has been a topic of investigation with a focus on the outcomes of the implementation of these tools. VR-assisted language learning (VRALL) provides “sensory-rich environments, allowing [learners] to experience telepresence (i.e., the feeling of ‘being there’ in the target language country)” (Kaplan-Rakowski & Wojdyski, 2018, p. 124). This immersive modality possesses the capacity to augment learner motivation and engagement, particularly in situations where exposure to authentic linguistic environments is constrained. One study that investigated the application of VR in the field of language learning found that VR facilitates learning experiences while eliminating geographical barriers for L2 learners (Kaplan-Rakowski & Wojdyski, 2018). This study also found that 82% of participants wanted to keep studying a language after using VR. Similarly, a systematic review of immersive VR applications in higher education found improved educational outcomes due to the integration of authentic tasks, real-time feedback mechanisms, and interactive simulations (Radianti et al., 2020).

In terms of the effects of VR on L2 vocabulary acquisition, there are relatively few studies despite the growing interest in the field (Agurto-Cabrera & Guevara-Vizcaíno, 2023; Moreno Martínez & Galván Malagón, 2020; Valero Franco & Berns, 2023). However, Legault et al. (2019) found that immersive VR can be an effective tool for L2 vocabulary acquisition as VR scenarios promote contextualized vocabulary use and can improve memory retention through multisensory experiences. Additionally, providing these tools to learners allows them to use their motor skills while learning an L2, which is again achieved because of the replicas of real-world scenarios where learners can engage with vocabulary in meaningful ways (Legault et al., 2019).

Other research investigated the advantages and limitations of using VR in L2 education. Advantages included having an authentic, real-life environment and the ability to engage with multiple senses leading to higher motivation, retention, and achievement (Klimova, 2021). VR supports diverse learning styles, fosters creativity, and boosts self-confidence while reducing anxiety and encouraging active participation and the development of learner autonomy. However, there are some limitations to consider, such as the high cost of software and teachers and students' potentially limited tech skills (Klimova, 2021). The success of VR applications relies on carefully designed content that follows cognitive and instructional design principles (Mayer, 2014). The challenge for educators and programmers, then, lies in the creation of VR experiences that integrate interactivity, immersion, and educational efficacy while ensuring accessibility and cost-effectiveness to promote wider adoption.

2.2.1.1 Grounded Cognition and Vocabulary Acquisition

VR provides not only an immersive environment for language learners but also sensorimotor experiences. This combination can be explained by the framework of grounded cognition, which emphasizes the important role of sensorimotor experiences in the language acquisition process while highlighting the importance of contextualized and experiential approaches in vocabulary acquisition. Barsalou (2008) suggested that cognitive functions, including language comprehension, are fundamentally intertwined with the brain's modal systems dedicated to perception, action, and introspection. This notion indicates that lexical items are not simply abstract representations. Instead, they are deeply rooted in sensorimotor experiences, reinforcing the importance of contextualized and experiential learning approaches in vocabulary acquisition. As individuals interact with their environment, the sensorimotor states linked to these experiences are encoded by memory systems for future representational use

(Barsalou, 2023). Therefore, grounded learning methodologies emphasize the importance of contextual and experiential vocabulary acquisition, where lexical items are learned through active participation in authentic real-world contexts rather than through passive retention. The multimodal characteristics of this learning approach ensure that vocabulary components are associated with rich, interrelated representations that include visual, auditory, tactile, and emotional dimensions and can also be closely connected to cultural factors, as vocabulary acquisition is influenced by culturally specific situated activities and linguistic expressions (Barsalou, 2023). This theory supports the integration of VR in L2 education.

2.2.2 AI and Language Education

In addition to VR, AI is another technological tool that can be used in education. AI is defined “as the branch of computer science dedicated to creating systems capable of performing tasks that typically require human intelligence” (Macinska & Vinkler, 2021, p. 4). Generative AI (GenAI) is characterized by its ability to generate new content that has not existed previously. As such, GenAI can create output that may appear original, and one example of this is chatbots that can mimic human conversations (Macinska & Vinkler, 2021).

Chatbots are software applications that simulate human conversation by asking and answering questions via text or audio. They are not new as they have existed since the 1960s, and they have been integrated into different messaging apps for years (Son et al., 2023). In the context of L2 learning, chatbots have been programmed to engage in conversations with the learner to perform different tasks. Research has demonstrated that the use of chatbots increases the output of the learner and also enhances their speaking performance (Son et al., 2023).

AI is revolutionizing language education by offering new ways to develop different linguistic skills. In their systematic review, Zhu & Wang (2025) highlighted different ways in which AI tools can be used in the process of language learning, including AI-powered writing evaluation systems that provide feedback on written work as well as other tools like ChatGPT. AI-driven avatars are also helping learners improve their pronunciation and speaking abilities. AI-enhanced language learning experiences have the potential to address multiple dimensions of language learning and proficiency development; however, GenAI was only released in an open access format in 2022, and much more research is needed to understand its full potential.

2.2.3 VR and AI in LSP

As this study is focused on vocabulary acquisition in LSP and integrates VR and AI, this next section presents relevant, albeit limited, research in this area. Li et al. (2022) studied the effects of an immersive VR experience on vocabulary acquisition in English for Geography. Through the interaction with multimodal and contextualized scenarios, such as the hydrologic cycle, researchers found that VR can improve incidental vocabulary learning and learner engagement as the participants were able to activate memory and comprehension due to the immersive environments (Li et al., 2022). Additionally, these environments supported cognitive, behavioral, and social participation, which are essential for mastering specialized terminologies (Li et al., 2022). Similarly, Park (2022) studied vocabulary learning in a VR-enhanced real-world environment called the Digital Kitchen. This study combined TBLT with multimodality to engage learners' senses (i.e., touch, smell, and taste), which are often absent in traditional or virtual settings. Results showed that learners in the Digital Kitchen achieved higher scores in vocabulary retention compared to those in traditional classroom settings. Park (2022) highlighted

the importance of integrating physical and sensory experiences into language learning to enhance memory and language acquisition.

Miller De Rutté (2024a, 2024b) investigated the role of VR in medical Spanish instruction. Undergraduate students enrolled in a medical Spanish course engaged in immersive VR simulations designed to replicate real-world professional interactions, allowing them to perform medical tasks in Spanish. Miller De Rutté (2024b) found increases in students' motivations levels after using VR as students expressed a heightened ability to visualize themselves successfully performing professional tasks. Additionally, students emphasized specific skill development, particularly in conversation, listening, pronunciation, and critical thinking (Miller De Rutté, 2024a). The immersive nature of VR was found to deepen students' self-reported levels of engagement and focus while amplifying course topics related to cultural awareness, pragmatics, and applied medical vocabulary. Participants expressed that the simulations created a more authentic and interactive learning environment, reinforcing their ability to connect language skills with professional application (Miller De Rutté, 2024a).

2.2.4 VR and AI in the Veterinary Field

Similarly to the use of technology in language education, VR and AI have also been implemented in the veterinary field as they offer new solutions to improve education, diagnostics, and administrative efficiency (Appleby & Basran, 2022; Sobkowich, 2025). Specifically, VR enables veterinary students and professionals to practice surgical techniques and diagnostic procedures in a controlled, risk-free environment (Xu et al., 2023). These immersive simulations not only improve skill acquisition but also build confidence, preparing practitioners for real-life scenarios (Appleby & Basran, 2022; Sobkowich, 2025). VR is also used in anatomy education, where students can explore 3D models of pet's anatomy derived from real

cases, and this approach has been shown to improve interactive learning while linking anatomical studies to clinical relevance (Aghapour & Bockstahler, 2022). Regarding the use of AI in veterinary medicine, Sobkowich (2025) discussed the variety of ways in which AI is being used in the field. Examples include automating workflow for scheduling appointments, predicting surgery complications, using adaptive testing systems, or extracting data from patient records. In the education field, Sobkowich (2025) states that AI is being used with platforms like Khanmigo to provide adaptive and on-demand support for students, while intelligent learning management systems adjust lesson pacing and sequencing based on individual progress. Another tool that is being used is AI-driven chatbots to recreate clinical scenarios or practice client communication. However, despite these technological advancements, the application of VR and AI combined specifically for communication in veterinary contexts remains largely unexplored.

2.3 Vocabulary Acquisition in Language Learning

Vocabulary acquisition serves as a fundamental function in the process of L2 learning with a direct influence on communicative competence and language proficiency. Vocabulary is part of the foundation for developing the four language skills (i.e., listening, reading, writing, and speaking). There are complex learning processes involved in vocabulary acquisition and retention, and retention processes are interconnected with the acquisition and transfer phases of the learning process in addition to memory and cognitive processes (Houston, 2001 as cited in Sanatullova-Allison, 2014). Sanatullova-Allison, (2014) explained that

once an item is perceived, it enters primary memory (PM) with short-term storage.

Rehearsal is necessary for the item to remain in PM and, if rehearsal is long enough, the item may enter secondary memory (SM), which is long-term storage [...] Long-term

memory is made of declarative and procedural knowledge: the former is the knowledge about facts and the latter is the knowledge about how to perform tasks (p. 2).

Different language learning and teaching principles have been implemented to promote L2 vocabulary acquisition and memory retention (Nation, 2022). For example, spaced repetition learning, where review intervals are lengthened after successful recall and shortened after mistakes, has been found to significantly improve long-term retention compared to studying in one continuous session (Sanatullova-Allison, 2014). Other research has shown that engaging in active recall, defined as a learning technique that involves actively retrieving information from memory rather than passively reviewing it (Pham et al., 2016), significantly improves memory retention, promotes deeper cognitive engagement, and leads to better learning outcomes unlike strategies such as re-reading or passive review (Karpicke & Roediger III, 2008). Similarly, in terms of instructional techniques to teach vocabulary, research has found that intentional vocabulary teaching generally results in long-term retention as well as proficiency gains in the L2 (Schmitt & Schmitt, 2020). Other research has found that learners who actively participated in tasks outperformed those who did not engage in tasks (Joe 1998, as cited by Sanatullova-Allison, 2014), which emphasizes the importance of task-based teaching approaches. Word frequency, how words are presented, what students do with the words, and what language is used when explaining vocabulary are other key considerations in instruction (Borawski, 2019). Presenting vocabulary in terms of frequency based on a spoken L2 corpus and using those words in a small and morphological word list has been found to improve the acquisition and retention of L2 vocabulary (Borawski, 2019).

Therefore, vocabulary knowledge is a construct that includes form, meaning, and use, and knowing a word is not only recognizing or defining a word but using it in the correct context.

Mastering all of these aspects enables a speaker to use the word with native-like fluency and accuracy (Schmitt & Meara, 1997). Previous research indicates four key aspects inherent to vocabulary knowledge, which include vocabulary size (the number of words known), collocational competence (the learner's ability to recognize and use word combinations that naturally occur together), receptive vocabulary knowledge (the ability to recognize words), and productive vocabulary knowledge (the ability to use words appropriately) (Masrai, 2023; Nation, 2022). Apart from these core dimensions, vocabulary knowledge is also influenced by depth of word understanding, including semantic associations and morphological awareness (Mcbride-Chang et al., 2008; Samaraweera, 2025). The relationship between these elements ensures that learners not only recognize and produce words but also interpret subtle meanings and contextual relevance. This process enhances their ability to communicate effectively in the target language.

2.3.1 Vocabulary Assessments

Evaluating L2 vocabulary comprehension is an essential element in language proficiency assessment. Over time, scholars have developed a variety of instruments to evaluate both the breadth and depth of learners' vocabulary knowledge with different tools measuring different vocabulary aspects. These instruments include traditional word-list exams as well as more advanced assessments based on contextual vocabulary use that evaluate the learner's capability to identify and use words within specific contexts. Some of the most common types of vocabulary tests are receptive vocabulary tests, productive vocabulary tests, and computer-adaptive testing.

Initial research in vocabulary assessment focused on measuring receptive vocabulary through breadth, which refers to the number of words learners know, with common tools like the Vocabulary Size Test (Nation, 1990; Nation & Beglar, 2007). However, it was later recognized

that vocabulary depth, or how well learners understand word meanings, associations, and usage, is also important to assess (Pignot-Shahov, 2012; Sun et al., 2023). Productive vocabulary tests assess the learner's ability to actively use words in different contexts, such as translation (Fitzpatrick, 2007). One example of this type of test is the Productive Levels Test (Laufer & Nation, 1999), which evaluates vocabulary knowledge based on word-frequency bands. For this test, participants reorganize letters to create a word, and there are a possible total of eight words across five levels (Fitzpatrick, 2007). Computer-adaptive testing, such as the Computer Adaptive Test of Size and Strength (CATSS), assesses both breadth and depth of vocabulary knowledge by measuring active recall, passive recall, active recognition, and passive recognition while adapting dynamically to the learner's proficiency level (Pignot-Shahov, 2012). However, the most widely used method is the receptive vocabulary test, which measures learners' ability to recognize words and is the type of test used in this study. This type of test does not require the learner to produce language, which makes it accessible to a wide range of proficiency levels (Amenta et al., 2020).

As mentioned, a common receptive vocabulary assessment is the Vocabulary Levels Test (Nation, 1990). It encompasses various levels of difficulty, from basic to more advanced. The assessment requires learners to match words with their corresponding meanings, and each level presents a distinct set of words that align with frequency bands in the English language (e.g., the 1,000 most frequently encountered words, the 2,000 most frequently encountered words, etc.). The Vocabulary Size Test, on the other hand, was created by Nation & Beglar (2007) to assess learners' vocabulary size. This test consists of a series of word recognition tasks designed to evaluate the number of words a learner can comprehend. The items are sourced from various

frequency bands in the English language, and the test provides an estimation of the total number of words with which a learner is familiar across the different frequency bands.

Another widely known assessment is the LexTALE test, introduced by Lemhöfer & Broersma (2012), which evaluates the receptive vocabulary size of advanced English learners. The LexTALE uses a lexical decision task format, requiring learners to decide whether the presented words are real English words or invented terms. The assessment is designed to be both quick and reliable and incorporates a scoring system that determines a learner's vocabulary size based on their performance. The LexTALE can differentiate between varying levels of language proficiency and is a validated measure. Building upon the success of the LexTALE, other researchers have created similar instruments for other languages. Izura et al. (2014) developed the Lextale-Esp, a test adapted for the Spanish language. The test consists of 60 items (40 real words and 20 non-words) and can differentiate across Spanish language proficiency levels, including at the higher ends of the proficiency scale. Amenta et al. (2020) developed the LexITA to measure receptive vocabulary size in Italian and has been validated as an effective assessment for evaluating vocabulary knowledge across a spectrum of proficiency levels, as its equivalents in Spanish and English.

These three tools, the LexTALE, Lextale-Esp, and LexITA are distinguished from others by their efficiency and accuracy. They are structured for quick administration, generally requiring five to ten minutes to complete, while also enabling objective scoring through a simple yes/no format. This study built on the strengths of these existing assessments, as it followed their methodology. However, it further adapted their work to create a specialized vocabulary test to address the needs of veterinary Spanish learners.

2.4 Research Questions

The purpose of this study was to understand the effects of the use of VR and AI on Spanish vocabulary acquisition and retention in a course on Spanish for Veterinarians. The following research questions (RQs) guided the study:

RQ1: Does the use of a VR and AI platform in the Spanish for Veterinarians classroom affect vocabulary acquisition in L2 learners?

RQ2: Does the use of explicit versus implicit vocabulary learning activities in a VR and AI environment influence learners' vocabulary acquisition and retention?

CHAPTER 3: THE DESIGN AND VALIDATION OF THE LEXVET-ESP TEST

3.1 Introduction to Project 1

The overarching goal of this thesis was to understand technology's influence on vocabulary acquisition and retention in students enrolled in courses on Spanish for Veterinarians. However, there is no validated vocabulary assessment to measure vocabulary acquisition in this area. Therefore, the main objective of this first part of the research was the creation of a receptive vocabulary acquisition test to serve as a data collection measure in the subsequent chapter. The LexVet-Esp test, as this new test is now called, followed the same procedures from previous research to develop and validate the new test (Amenta et al., 2020; Izura et al., 2014; Lemhöfer & Broersma, 2012), which is described in the following subsections. The specific research questions guiding this first project were as follows:

RQ1: How accurately does the LexVet-Esp predict Spanish vocabulary knowledge and proficiency?

RQ2: How well does the LexVet-Esp differentiate between native and non-native Spanish speakers?

3.2 Methodology

3.2.1 Materials

Ninety words were extracted from a spoken corpus on Spanish for Veterinarians (Zeller et al., forthcoming). The corpus was based on veterinary medicine observations from four different locations, Colorado, Texas, Mexico, and Colombia, and included observations in both Spanish and English. For this project, only the Spanish interactions were analyzed. These

interactions occurred between different personnel in a veterinary clinic (e.g., veterinary technician, veterinarian, etc.) and the Spanish-speaking client. The selection of the 90-word items was based on frequency distributions as described in previous studies (Amenta et al., 2020; Izura et al., 2014; Lemhöfer & Broersma, 2012). Words spanned from exceedingly high frequency words (e.g., *perro*/dog or *doctor*/doctor), which are anticipated to be familiar to most Spanish speakers, to very low frequency words, (e.g., *férula*/splint or *sarro*/tartar), which are likely to be recognized exclusively by highly proficient primary language speakers.

The characteristics of the word items, including their distribution, length, and frequency per million words, are presented in Table 1. Overall, 26 words had a frequency of less than one occurrence per million words, 23 had a frequency of one to five occurrences per million, 14 had a frequency of 6 to 10 occurrences per million, 17 had a frequency of 11 to 20 occurrences per million, eight had a frequency of 21 to 100 per million, and two words (*comer*/to eat and *mirar*/to look) had a frequency greater than 100 per million. Most of the words were nouns (n=50), and there were equal numbers of verbs (n=20) and adjectives (n=20).

Table 1

Characteristics of the word items included in the initial test with 90 words

Distribution	Length (letters)	Frequency per million
min	3	22.22
1 st quartile	6	111.11
median	8	144.44
3 rd quartile	9.75	88.88
max	15	33.33

mean	8.13	144.44
SD	2.58	677.77

Next, a compilation of 90 non-words was generated. These non-words were pseudowords, but they followed similar patterns to real words. For example, verbs in Spanish end in -ar, -er, or -ir, and non-word verbs followed that same pattern (i.e., real word - *comer*; non-word - *riñonar*). ChatGPT 3.5 was used to generate non-word items after it was given the authentic word list and was prompted to create 90 non-words, including nouns, verbs, and adjectives, similar to the real word list. It was also prompted to use different parameters, such as including suffixes analogous to various grammatical categories (e.g., adding *mente* for an adverb) or verbs with Spanish characteristics (e.g., ending in -ar, -er, -ir). ChatGPT 3.5 generated non-word parts of speech, including nouns (*toráciz, cachorro, parasíticoz*), adjectives (*pulgario, abdominado, anestesiático*), and verbs (*autolimitantear, alérgizer, brinquir*).

After the creation of these 90 non-word items, they were reviewed to verify that the words did not exist. To do this, two dictionaries were consulted – the *Real Academia Española* (2014) and the *Real Academia Nacional de Medicina* (2012). Non-words were searched in both dictionaries to verify that they did not exist in either of them. If they did exist, a new non-word was created for which it could be substituted. There were only ten words that ChatGPT 3.5 generated as non-words but were, in fact, real words. They included *ótico* (otic), *lipémico* (lipemic), *intratorácico* (intrathoracic), *cremar* (to cremate), *sondeo* (probing), *remisivo* (remissive), *venenoso* (poisonous), *blandir* (brandish), and *esterilizante* (sterilizing). Table 2 shows the distribution, length, and frequency per million words of the selected non-words. Note

that the frequency per million words is listed as N/A due to the fact that these are non-words, and so, frequency per million words cannot be calculated.

Table 2

Characteristics of the non-words items included in the initial test with 90 non-words

Distribution	Length (letters)	Frequency per million
min	3	N/A
1 st quartile	6	N/A
median	8	N/A
3 rd quartile	9.75	N/A
max	15	N/A
mean	8.13	N/A
SD	2.58	N/A

3.2.2 Procedure and Participants

To determine which real words and non-words were to be included in the vocabulary assessment, the list of all words was distributed to students at different proficiency levels in the form of a written paper test. The test (Appendix A) was administered in person so that the participants did not have access to the internet nor could they consult a dictionary or online translator to verify whether a word existed or not. In the first part of the test, participants answered a short questionnaire where they indicated the level of the class they were taking, their age, race, and what their primary language was. Then, participants were given a written explanation of the test with an example of four common words (*sí* - yes; *sacapuntas* - pencil

sharpener; *bien* - good; *casa* - house) so that they could better understand how the test functioned (Amenta et al., 2020). The instructions for this first part of the test were provided in English.

This test was given to students enrolled at a large western university in the U.S. Participants were divided into two main groups. The first group consisted of primary Spanish speakers (n=23), and the second group was both heritage Spanish speakers and Spanish L2 speakers (n=180). The second group was further divided according to the level of the class in which the participants were enrolled, following previous research. There were 49 participants at the 100-level, 35 at the 200-level, 63 at the 300-level, 30 at the 400-level, and three at the 500-level. Participants were then given the 180 items (90 real words and 90 non-words) that were randomly ordered, and they were asked to identify the real words on the list. All participants received the same order of real words and non-words. Participants were not given a specific time limit to complete the test; however, no participant took more than 15 minutes to complete it.

3.2.3 Data Analysis

The first part of the assessment collected demographic data. This data was analyzed using descriptive statistics and was used to separate the participants into two main groups (native speakers vs. heritage and L2 speakers) as well as across the different proficiency levels. The vocabulary test was then analyzed using a series of statistical tests to determine which 60 real words and 30 non-words were to be included in the final version. First, following previous research (Lemhöfer & Broersma, 2012; Brysbaert, 2013), the tests were scored using the following formula:

$$\text{Score} = N_{\text{yes to words}} - 2 * N_{\text{yes to non-words}}$$

This formula was used as it penalized participants for random guessing behavior, meaning that if a participant answered arbitrarily (i.e., affirmatively responding to half of the real words and half of the non-words), they would obtain a score close to zero. This formula also allows for participants to earn a negative score if they incorrectly categorized non-words as real words.

Then, point-biserial correlations and Item Response Theory (IRT) were conducted. Point-biserial correlation coefficients range from -1.0 to +1.0. For the purposes of this study, participants with a positive correlation exhibited superior performance on the overall test while also excelling on the specific item in question. However, a negative correlation implied that the test was not successfully measuring the intended outcome, as it suggested that proficient participants were performing less favorably on the specific item compared to their lower proficiency counterparts. An IRT analysis was also conducted to determine if items on the assessment were accurately measuring the proficiency of the participants. This type of analysis provided information related to both the difficulty level of the item and its discriminatory power. Discriminatory power refers to the steepness of the item response curve, transitioning from “not known” at the lower end of the ability spectrum to “known” at the upper end.

Once the most effective test items were selected, test performance across groups of participants was compared using an Analysis of Variance (ANOVA) to determine if the test could discriminate across proficiency levels. The R package *ltm* (Rizopoulos, 2006) was utilized for all analyses in this study.

3.3 Results

Participants' (n=203) ages ranged from 18-70 with a mean age of 21.65. See Table 3 for mean, standard deviation (SD), and age range by class level. The majority of participants' self-

identified as White (n=142; 69.95%), which was followed by Hispanic/Latino (n= 44; 21.67%), other (n = 11; 5.42%), or Black/African American (n = 6; 2.96%). When asked about their primary language, the majority replied that English was their L1 (n= 178; 87.68%), while a smaller portion indicated that Spanish was their first language (n = 25; 12.32%).

Table 3

Distribution of age according to the level of the class

Group	Mean	SD	Range
100	21.51	7.53	47 (65-18)
200	18.94	8.50	52 (70-18)
300	20.25	1.38	7 (25-18)
400	18.74	9.75	51 (70-20)
500	25.5	3.62	8 (30-22)
Native speakers	29.8	8.81	22 (43-21)

3.3.1 Point-biserial Correlation

To select the items to be included for the final test, the quality of both the real words and non-words was analyzed by calculating the point-biserial correlation between the responses to each item and the cumulative scores of the participants. All items had a positive correlation (ranging from $r = 0.08$ for the word *hez/feces* to $r = .68$ for the word *glándula/gland*). Since there were only positive correlations, no items could be eliminated, and an IRT needed to be conducted next.

3.3.2 Item Response Theory

The results from the IRT analysis produced charts with difficulty of the word item along the x-axis and probability of correctly answering the item along the y-axis. This information then allowed for the selection of items that spanned a broad spectrum of difficulty using discriminatory power, which was measured by the steepness of the response curve at its midpoint. A steeper curve indicated greater discriminatory power of the item. Figures 1 and 2 show the IRT charts for real words and non-words.

Figure 1

IRT of real words

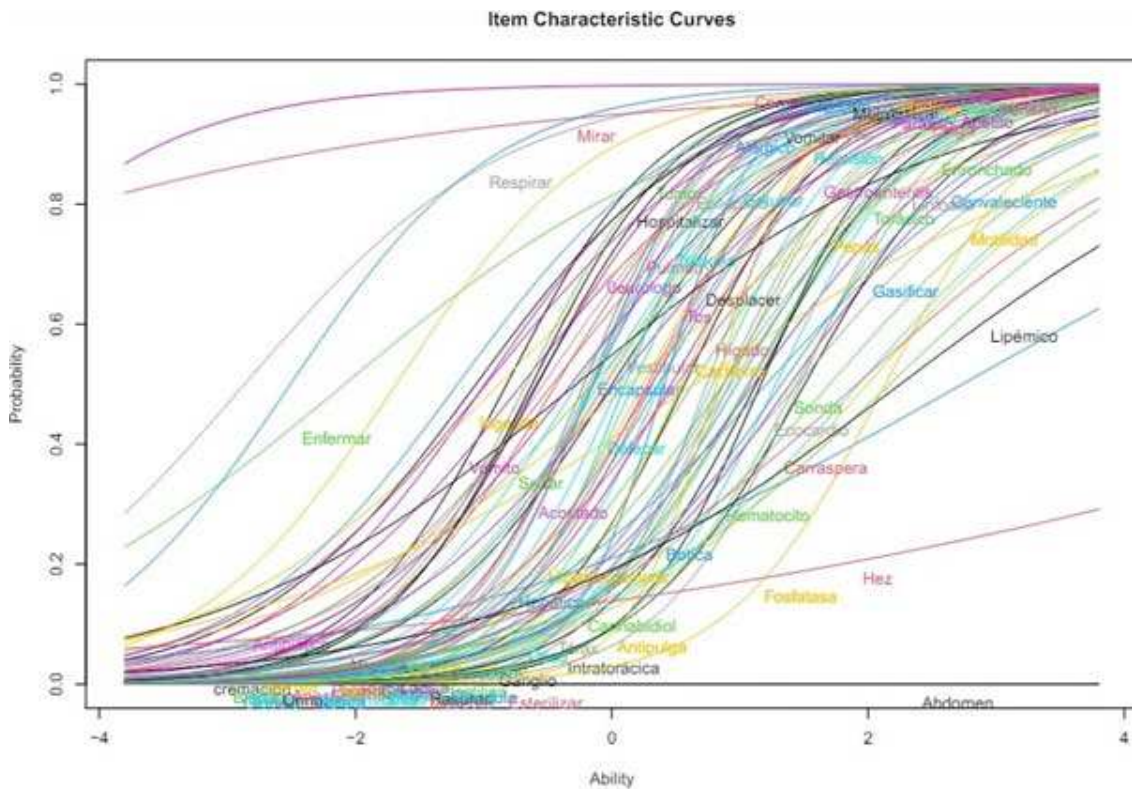


Table 4 presents a selection of sample words, where the intercept represents the IRT item difficulty, and z1 indicates the discriminative power of each word relative to the others.

Table 4

Discriminative Power and IRT Scores for Selected Items

Item	Intercept	z1
<i>Cremación</i> (cremation)	1,338,778,108	16,326,639
<i>Irritar</i> (to irritate)	0.508833260	11,035,564
<i>Remisar</i> (non-word)	-4,156,332,499	0.1710680
<i>Enroncharar</i> (non-word)	-4,156,332,499	0.1710680

Based on these results, the final word list was determined, and descriptive statistics can be seen in Table 5. The table presents the distribution of word and non-word lengths. Word length ranged from three to 14 letters, with a median of nine, while non-word length spanned from seven to 15 letters, with a median of 11. The frequency per million varied across words, averaging 16,821.35, whereas non-words do not have frequency values assigned.

Table 5

Descriptive statistics of the final words

Distribution	Words		Non-words	
	Length (letters)	Frequency per million	Length (letters)	Frequency per million
min	3	2,320.19	7	N/A
1 st quartile	7	2,320.19	9	N/A
median	9	8,120.65	11	N/A

3 rd quartile	11	23,201.86	12	N/A
max	14	102,088.17	15	N/A
mean	8.98	16,821.35	10.74	N/A
SD	2.94	21,472.48	2.53	N/A

3.3.3 Comparisons across proficiency levels

The formula for scoring the final test as described previously was used to calculate participants' scores. Possible scores ranged from -180 (if a participant selected all non-words as real and all real words as non-words) to 90 (if a participant accurately identified all real words and refrained from selecting any non-words as real). Table 6 shows the groups of participants by level along with their mean score, standard deviation, and the range of test scores.

Table 6

Test scores by proficiency group

Group	Mean score	SD	Range
100	-5.59	16.52	63 (-44 – 19)
200	1.97	20.66	114 (-85 – 29)
300	15.28	17.59	96 (-30 – 66)
400	18.06	18.36	71 (-25 – 46)
500	47.67	25.10	54 (18 – 72)
Native speakers	66.20	11.63	28 (52 – 80)

A one-way ANOVA was conducted to determine the difference in mean scores across the six proficiency groups. Overall, there was a statistically significant difference among the groups ($F(5, 197) = 25.852, p < 0.001$). Tukey post hoc analyses revealed evidence of a significant difference between all levels except 100 vs. 200 ($p = 0.4089$), 300 vs. 400 ($p = 0.9759$) and 500 vs. native speakers ($p = 0.5451$).

3.4 Discussion

The objective of the initial investigation was to create and validate an instrument to test the acquisition of receptive vocabulary relevant to Spanish for Veterinarians. To do so, test items that possess the capability to effectively differentiate across proficiency levels were statistically determined. Unlike general vocabulary tests, such as the LexITA (Amenta et al., 2020) or the Lextale-Esp (Izura et al., 2014) which measure general vocabulary, this instrument focuses on technical terms for veterinary professionals. This section discusses the effectiveness of the LexVet-Esp in assessing specialized vocabulary knowledge by comparing its results to existing assessments, evaluating the strengths and limitations of the statistical analyses, and determining its ability to differentiate proficiency levels among participants. Finally, limitations and future directions are presented.

3.4.1 Comparison of the LexVet-Esp with LexITA and Lextale-Esp

While the LexITA (Amenta et al., 2020) and Lextale-Esp (Izura et al., 2014) evaluate general receptive vocabulary proficiency, the LexVet-Esp has been specifically developed to assess field-specific vocabulary knowledge within the area of Spanish for Veterinarians. Although all three assessments use a combination of real words and non-words to measure vocabulary recognition, the LexVet-Esp prioritizes technical terminology in veterinary consultations, which ensures its relevance for professionals of the field. To prioritize this

specialized terminology, the selection of real word items was guided by a spoken corpus of veterinary interactions in Spanish.

The methodology for the selection of non-words was likewise informed by the same approaches used in the LexITA and Lextale-Esp instead employed artificial intelligence alongside manual validation to ensure that the non-words resembled real words in form and structure while lacking semantic value (Amenta et al., 2020; Izura et al., 2014). This step is critical in mitigating guessing behavior, as participants are forced to identify between real words and non-words, which is a fundamental aspect of assessing receptive vocabulary (Ha, 2021; Masrai, 2023). The incorporation of non-words assured that participants were required to possess actual lexical knowledge to distinguish real words from non-words, thereby enhancing the reliability of the assessment.

3.4.2 Addressing the Limitations of Point-Biserial Correlations with IRT

All test items displayed positive point-biserial correlations, indicating that participants who performed well overall were more likely to correctly identify individual items. However, while this confirmed a basic relationship between item difficulty and participant ability, the analysis did not provide additional information to differentiate item performance across varying proficiency levels. This limitation aligns with prior observations in the LexITA (Amenta et al., 2020) and Lextale-Esp (Izura et al., 2014), where point-biserial analysis was useful for initial evaluation but insufficient for deeper item validation. Given the need for a more precise assessment of individual items, IRT was employed to provide a more comprehensive analysis of item difficulty and discrimination across proficiency levels. The IRT analysis demonstrated the assessment's capacity to distinguish across learners of varying proficiency levels, with both simpler and more challenging items effectively differentiating individuals with lower proficiency

from those with higher proficiency. The broad variety of word difficulties incorporated into the final assessment was necessary to guarantee that the evaluation was sensitive to the full spectrum of language proficiency levels.

3.4.3 Differentiation of Proficiency Levels

The LexVet-Esp scores demonstrated significant differences across proficiency levels. Higher-proficiency groups, particularly those at the 500-level and native speakers, tended to perform better in correctly identifying vocabulary items. This finding partially supports the hypothesis that vocabulary acquisition is progressive (Nation, 2022), with learners expanding their lexicon to include increasingly complex and less commonly used terminology as they advance in their studies (Borawski, 2019). Lower-proficiency learners showed limitations in identifying specialized terminology. This trend aligns with the Lextale-Esp findings, where participants at lower proficiency levels struggled to identify infrequent vocabulary items (Izura et al., 2014). The data in the current study reflect similar patterns, reinforcing the notion that specialized vocabulary knowledge is not solely a product of exposure but also an indicator of overall linguistic proficiency (Masrai, 2023). Additionally, and consistent with previous studies (Amenta et al., 2020; Izura et al., 2014), participants who frequently selected non-words tended to score lower overall. This result aligns with prior research suggesting that non-word recognition can serve as a general indicator of language proficiency (Roche & Harrington, 2013). Overall, the results indicate the test's ability to differentiate learners ranging from novices to advanced speakers. This finding underscores the instrument's potential for assessing specialized language acquisition with a level of precision comparable to established vocabulary assessments.

3.4.4 Limitations

Firstly, while the sample size was considered sufficient for analyses, the reliance on a single data collection site limits the extent to which the findings can be generalized. With the incorporation of participants from various geographical areas and/or backgrounds, especially those with different levels of exposure to the Spanish language, the validity of the results would be significantly enhanced. This consideration is especially relevant when reflecting on previous assessments, such as LexITA and Lextale-Esp, which featured participants from multiple universities and diverse linguistic backgrounds, strengthening the credibility of their conclusions (Amenta et al., 2020; Izura et al., 2014). Additionally, the small number of participants at the 500 level (n=3) represents a clear limitation as it limits the insights of higher proficiency learners. Another limitation is that the proficiency level was measured by course enrollment level, which may not be the most accurate measure of language proficiency. Future research should incorporate a standardized proficiency test to determine the level of participants.

CHAPTER 4: MEASURING VOCABULARY ACQUISITION WITH THE LEXVET-ESP IN A VIRTUAL REALITY ENVIRONMENT

4.1 Introduction to Project

The principal objective of Project 2 was to implement the LexVet-Esp in a Spanish for Veterinary Medicine course to test participants' vocabulary acquisition and retention before and after using a platform that combined VR and AI. The project also aimed to compare the results of two distinct instructional modalities for vocabulary acquisition: explicit versus implicit vocabulary exercises. Explicit vocabulary instruction involves direct explanation of word meanings, uses, and forms, followed by focused practice. In contrast, implicit instruction integrates vocabulary exposure within wider communicative activities, avoiding direct emphasis on the words. By contrasting these approaches, the project attempted to gain insights into the roles of explicit and implicit instruction in vocabulary acquisition in a VR/AI-enhanced educational platform. Information regarding the platform and activities will be detailed in the upcoming sections. The specific research questions guiding this project were as follows:

1. What is the impact of a VR/AI platform on the acquisition of receptive vocabulary in a specialized language course on Spanish for Veterinary Medicine?
2. How does the use of explicit versus implicit vocabulary activities impact students' vocabulary learning in a VR/AI-based context in a course on Spanish for Veterinary Medicine?
3. How effective is VR/AI-based vocabulary practice for long-term retention of specialized vocabulary in Spanish for Veterinary Medicine?

4.2 Methodology

4.2.1 Participants and Context

Participants in this project were students enrolled in the Spanish for Veterinary Medicine certificate in the DVM program at a large, western university in the U.S. Out of all the students taking the certificate, fifteen completed all phases of this project and were included in the study. Participation in this study was optional and did not impact students' final grades.

Participants were enrolled in the second two-credit course of the elective certificate. The course met weekly for a duration of 50 minutes, and students were expected to complete two hours of work outside of class time. The course followed a flipped classroom in which students were required to complete a series of online preparatory activities prior to each face-to-face class. These preparatory materials included tasks designed to develop reading comprehension, listening proficiency, vocabulary enhancement, and speaking skills and to familiarize students with essential thematic content and vocabulary pertinent to the in-person class. The face-to-face sessions prioritized active, communicative use of the language. During class sessions, students engaged in structured speaking exercises, discussions, and task-oriented learning activities that reinforced the material addressed in the preparatory phase.

4.2.2 Materials

A non-immersive VR/AI platform, MeTabi, was used for this project. MeTabi is an online language learning platform dedicated to enhancing students' linguistic proficiency within specified professional contexts that do not require headsets and spatial interaction, just the use of a computer or phone. Currently, it is available in over ten languages and uses Language Coaches, or characters that use AI, to help create a scaffolded learning process for students. This

framework permits the development of a variety of activities that focus on four essential skills: oral comprehension, written comprehension, oral production, and written production. As MeTabi is an AI-enhanced but screen-based system, learners benefit from immediate feedback generated by the AI system but do not experience real-time sensory immersion. For the purposes of this study, a virtual veterinary clinic was built in the MeTabi environment (Figures 3 and 4), and MeTabi was implemented into the Spanish for Veterinary Medicine course during the second half of the semester. Students used Metabi to complete their required outside of class work.

Figure 3

One of the consultation rooms in MeTabi

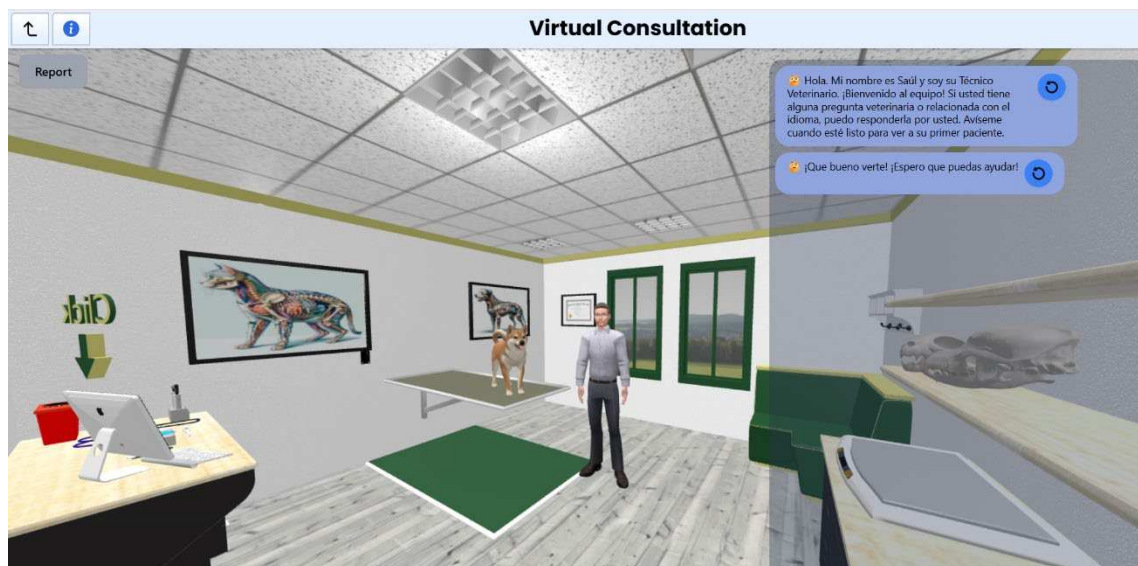


Figure 4

The clinic's laboratory with different diagnostic tools



From week nine to week twelve, students completed explicit vocabulary activities.

Examples can be seen in Figures 5-9. Figure 5 shows an activity in which participants watched a video talking about colors, textures, pet behaviors, etc., and images appeared alongside a list of words describing those images. At the end of the video, the participants completed a drag and drop activity where they moved each word into the correct category. Figure 6 shows a picture matching activity, and Figure 7 is a crossword puzzle. Figures 8 and 9 represent two parts of the same activity. For this activity, participants were presented with 18 flashcards, and each contained a picture and the corresponding word in Spanish. The first part of the activity required participants to translate the words into English, and in the second part, they were asked to complete a dialogue between the client and the doctor using the words from the flashcards (Figures 8 and 9).

Figure 5

Drag and drop activity

Organiza el siguiente vocabulario en la categoría correcta (drag and drop)

Colores de las heces

Texturas del excremento

Verde Naranja Líquido

Protector Rojo Blando

Sinónimos de heces

Temperamento de la mascota

Deposición Boflga Marrón oscuro

Caca Rojo pálido Jueguetón

Color de la sangre

Pegajoso Peligroso Con moco amarillo

2:38 2:38

Figure 6

Matching activity

Respiratory Symptoms - Vocabulary Activity (Part 1) Next Activity →

Arrastra las imágenes de la izquierda para emparejarlas con las frases correspondientes de la derecha.

Olores fuertes Mocos verdes Toser Olfato Estornudar

Moquear Respirar Secreciones oculares Congestión nasal

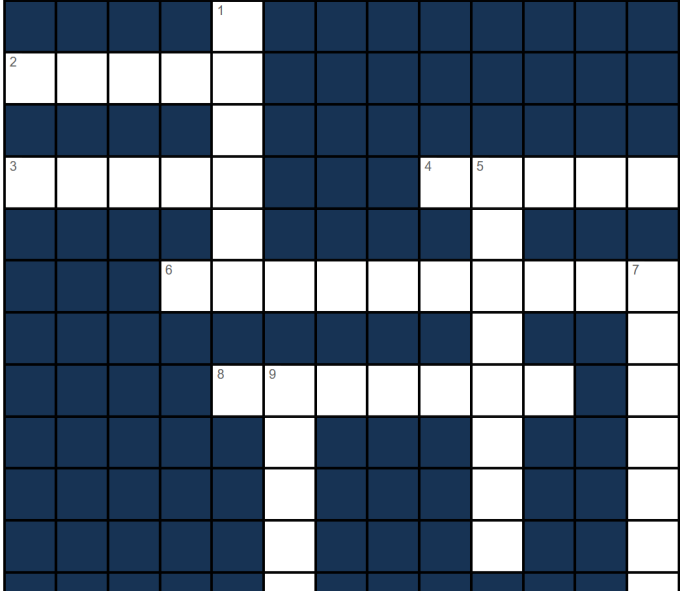
Check

Figure 7

Crossword puzzle

Respiratory Symptoms - Vocabulary Activity (Part 2)

¿Cuáles son las palabras? Organiza las letras para crear una palabra.



Across

- 2 Congestión anlas (5)
- 3 seort (5)
- 4 ocosm verdes (5)
- 6 teadsorunr (10)
- 8 arqomeu (7)

Down

- 1 eeorsl fuertes (6)
- 5 Secreciones soecura (8)
- 7 asreripr (8)
- 9 tofaol (6)

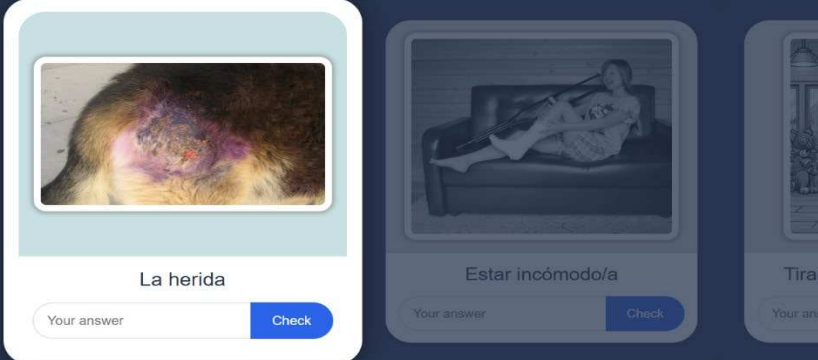
OneDrive - Personal
Uploading 936.0 KB of 1.1 MB at 0.0 KB/s. 1 file remaining

Figure 8

Part 1 of the flashcard activity

Primero vas a ver diferentes palabras del vocabulario relevante para esta semana. Vas a poder ver la palabra en español, además de una imagen y tienes que escribir su traducción en inglés.

1 / 18



La herida

Your answer

Estar incómodo/a

Your answer

Tirar

Your answer

Figure 9

Part 2 of the flashcard activity

Completa el diálogo entre el veterinario y su cliente con las palabras adecuadas.

Veterinario: Buenos días, ¿en qué puedo ayudarle hoy?

Cliente: Hola, doctor. Mi perro tiene una [] en la [] y él parece estar muy []. Además, no para de []

Veterinario: Entiendo, eso suena incómodo para él. ¿Ha notado algunas [] o irritación cerca de la herida?

Cliente: Sí, está un poco [] alrededor. Y también parece tener una roncha en la panza. Además, ahora también ha empezado a lamerse los [] .

Check

dedos
herida
irritada
rascarse
ronchas
incómodo
panza

Completa el diálogo entre el veterinario y su cliente con las palabras adecuadas.

Veterinario: Es posible que tenga una enfermedad llamada [] O también puede ser una [] . ¿Ha cambiado su dieta recientemente?

Cliente: No, no hemos cambiado su comida. Pero sí ha estado teniendo [] últimamente. Se está tirando muchos [] y no para de [] .

Veterinario: Los problemas gastrointestinales podrían estar relacionados con su herida en la [] . ¿Sus pedos [] muy fuerte?

Cliente: Sí, son bastante [] últimamente.

Veterinario: Voy a recetarle unos [] para tratar sus ronchas de la panza porque están muy [] y también para sus problemas gastrointestinales.

dermatitis
lamerse
antibióticos
huelen
pedos
apestosos
barriga
problemas gastrointestinales
inflamadas
alergia

During the remaining three weeks, there were implicit vocabulary exercises. Vocabulary was introduced through listening comprehension activities, oral activities with the Language Coaches, and pronunciation practice. No words were highlighted or explained to the students, and instead, the focus was on seeing how words were used in context. An example of this type of activity can be seen in Figure 10. In this example, participants watched a video after which they were asked to create five questions that they would ask their client in order to gather additional information about their pet. Another type of activity involved an AI Language Coach, as seen in Figure 11. For this activity, the participants were given a prompt that guided them to simulate a conversation in which they needed to use specific vocabulary that appeared in previous activities, even though those activities focused on grammar or listening comprehension. Another example activity, seen in Figure 12, included participants listening to three different clients, taking notes regarding the pet's health history, and then recommending a diagnostic procedure. This entire

activity required participants to utilize vocabulary seen in prior activities without explicit instruction.

Figure 10

Video with listening comprehension and writing questions

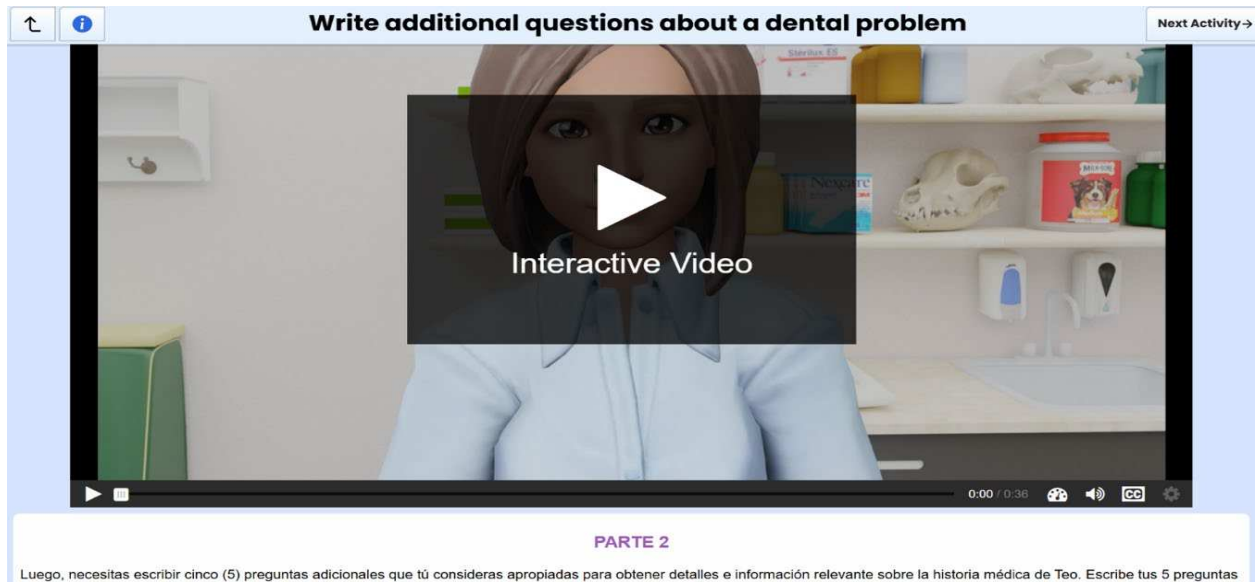


Figure 11

Language Coach activity

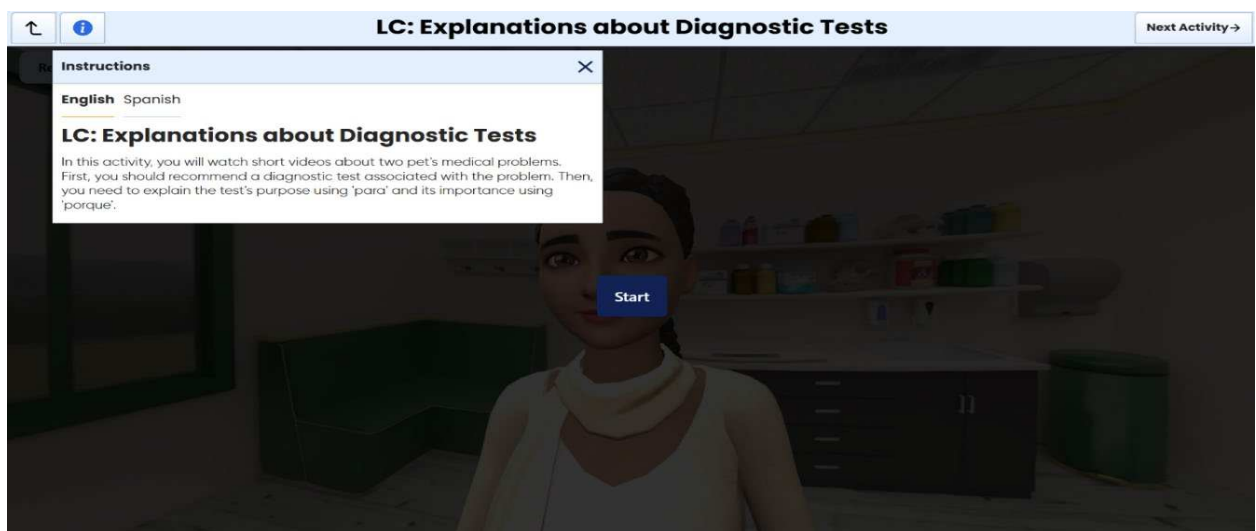
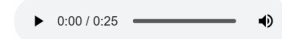


Figure 12

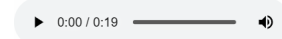
Listening and speaking activity

Escucha lo que dicen los siguientes 3 clientes y toma nota de algunos detalles de la historia médica en tu cuaderno o tableta.

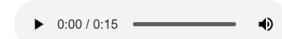
Cliente 1



Cliente 2



Cliente 3



> Haz clic aquí para ver las transcripciones de los audios de los clientes.

Usa la información de la historia médica de cada paciente para recomendar una prueba de diagnóstico y explicarle al/a la cliente/a **para qué** sirve la prueba y **por qué** es importante. Además, usa una frase al final para confirmar la comprensión del/de la cliente/a (*¿Me explico? ¿Es claro? ¿Todo es claro? ¿Me hago entender? ¿Fue claro todo lo que le expliqué? ¿Tiene sentido?*). Escoge dos (2) de los clientes y haz una grabación de audio o video y súbela a Canvas.

4.2.3 Procedure

The LexVet-Esp was implemented to evaluate students' vocabulary acquisition and was made up of two parts. The first part included a short questionnaire where students indicated their age, race/ethnicity, and their L1. This was followed by an explanation of the test with an example of common words so that they could better understand how the test functioned (Appendix B). The instructions for this part of the test were provided in English. Then, participants completed the LexVet-Esp by identifying real words from the list of 90 total words, 60 of which were real words and 30 were non-words. Participants were not given a specific time limit to complete the test; however, no participant took more than 15 minutes to complete the test.

This assessment was administered on three distinct occasions throughout the academic semester. Each time, there was a shuffling in the arrangement of the presented words. This modification was executed utilizing an online platform (<https://onlinetools.com/random/shuffle->

[words#tool](#)) and was done so that participants were unable to detect any patterns across the various versions of the test. On each testing occasion, participants received the same order of shuffled words.

The initial administration of the test occurred at the start of week nine of a sixteen-week semester and functioned as the pre-test since students had not yet started using the MeTabi platform. The second administration of the assessment was in week twelve. The distinction between weeks nine and twelve lies in the fact that, during those three weeks, the participants had engaged with MeTabi and had participated in explicit-only vocabulary exercises. The assessment was re-administered in week fifteen after participants had continued to use MeTabi but with implicit-only vocabulary exercises. This test administration allowed for the examination of potential changes between week nine and week fifteen, as well as the exploration of possible variations resulting from explicit versus implicit vocabulary exercises.

Then, three months after the semester ended, participants completed the same assessment for a fourth time. This final administration aimed to determine whether the students, after a summer break, retained any acquired vocabulary. Sixteen participants completed the first three tests. During the three-month break, one participant dropped out of the program, and their data was removed from the study. Therefore, the final number of participants who completed all four tests was 15.

4.2.4 Data Analysis

To analyze the data, the initial step involved substituting the participants' names with a unique and random identifier. Then, the responses from each participant were input into a spreadsheet and coded. The procedure for identifying the real words (a total of 60 words)

involved highlighting each real word with a distinct color. Then, all column headings corresponding to the real words were changed to a green color, while those pertaining to the non-words were changed to gray. This methodology was implemented to enable the conversion of the participants' responses into numerical values, facilitating the computation of the test outcomes. Then, participants' tests were scored, and the scoring process followed the procedure from the first study, which was based on previous research (Amenta et al., 2020; Izura et al., 2014; Lemhöfer & Broersma, 2012). On the test itself, participants indicated if words were real or not with a *Y* for "yes" and an *N* for "no." All instances of an *N* answer were converted to 0, as the failure to identify words neither detracts from nor enhances the overall score. Conversely, all instances of *Y* within the green columns, meaning that students correctly identified the real words, were assigned a score of 1. Lastly, all instances of *Y* in the gray columns, meaning that students incorrectly identified a non-word as real, were allocated a score of -2. A final score was calculated for each participant for each of the four tests. Descriptive statistics and repeated measures ANOVAs were used to analyze the data.

4.3 Results

The following section presents the results of the four tests conducted to evaluate participants' performance. The mean age of the participants was 25.67 years old. Most participants identified as White/Caucasian (n=13; 86.67%), while the rest identified as Hispanic/Latino (n=2; 13.33%). All participants reported their primary language as English. It is also important to note that prior to the start of this study, students had not engaged with any other VR/AI platform in the course.

Table 7 shows the mean scores of each of the four times participants completed the assessment. The first assessment had the lowest average score of 23.87, and the highest average

score, 27.87, occurred during the third assessment. The highest possible score that could have been achieved on this test was 60, and the lowest possible score was -120.

Table 7

Descriptive statistics of each test results

Test	Mean	SD	Range
1	23.87	5.68	21(34-13)
2	27.13	7.71	30 (41-11)
3	27.87	7.34	26 (42-16)
4	26.86	7.01	23 (41-18)

Then, each test was analyzed to determine which words participants identified as real words or non-words. For Test 1, only nine words were identified as real by all participants. Additionally, there were five non-words that all participants identified as not real, and there were two real words that none of the participants identified as real. A list of these words can be seen in Table 8. This same analysis process was followed for each test. For Test 2, 11 words were identified correctly as real by all participants. There were seven non-words that all participants identified as unreal, and there was one real word that none of the participants identified as real. A list of these words can be seen in Table 9.

Table 8

List of the words and non-word identified and not identified by the participants in Test 1

Real words correctly identified	Non-words correctly identified	Real words not identified
<i>Cachorro</i> (puppy)	<i>Comalir</i>	<i>Carraspera</i> (itchy throat)
<i>Defecar</i> (to defecate)	<i>Desplacante</i>	<i>Plastilina</i> (modeling clay)
<i>Diarrea</i> (diarrhea)	<i>Ecocardiante</i>	
<i>Mascota</i> (pet)	<i>Intratoracoide</i>	
<i>Orinar</i> (to urinate)	<i>Medularer</i>	
<i>Pulga</i> (flea)		
<i>Riñón</i> (kidney)		
<i>Tos</i> (cough)		
<i>Vacuna</i> (vaccine)		

Table 9

List of the words and non-word identified and not identified by the participants in Test 2

Real words correctly identified	Non-words correctly identified	Real words not identified
<i>Abdomen</i> (abdomen)	<i>Antibióticar</i>	<i>Carraspera</i> (itchy throat)
<i>Blando</i> (soft)	<i>Boticar</i>	
<i>Cachorro</i> (puppy)	<i>Carrasperante</i>	
<i>Defecar</i> (to defecate)	<i>Creatinar</i>	
<i>Diarrea</i> (diarrhea)	<i>Ecocardiante</i>	
<i>Esterilizar</i> (to spay/neuter)	<i>Intratoracoide</i>	
<i>Hepático</i> (hepatic)	<i>Medularer</i>	
<i>Orinar</i> (to urinate)		
<i>Pulga</i> (flea)		
<i>Tos</i> (cough)		
<i>Vacuna</i> (vaccine)		

On Test 3, all participants correctly identified ten real words and 11 non-words. There were three real words that none of the participants identified as real. A list of these words can be seen in Table 10. Finally, for Test 4, ten words were recognized as real by all participants. There were 15 non-words that all participants correctly classified as not real, and there were five real words that were not recognized as such by any of the participants. A list of these words is presented in Table 11.

Table 10

List of the words and non-word identified and not identified by the participants in Test 3

Real words correctly identified	Non-words correctly identified	Real words not identified
<i>Abdomen</i> (abdomen)	<i>Boticar</i>	<i>Carraspera</i> (itchy throat)
<i>Cachorro</i> (puppy)	<i>Comalir</i>	<i>Gasificar</i> (to gasify)
<i>Defecar</i> (to defecate)	<i>Creatinar</i>	<i>Plastilina</i> (modeling clay)
<i>Diarrea</i> (diarrhea)	<i>Desinflamantear</i>	
<i>Esterilizar</i> (to spay/neuter)	<i>Desplacante</i>	
<i>Gastroenteritis</i> (gastroenteritis)	<i>Diarreante</i>	
<i>Mascota</i> (pet)	<i>Intratoracoide</i>	
<i>Pulga</i> (flea)	<i>Medularer</i>	
<i>Tos</i> (cough)	<i>Remisar</i>	
<i>Vacuna</i> (vaccine)	<i>Sedarizar</i>	
	<i>Tumórico</i>	

Table 11

List of the words and non-word identified and not identified by the participants in Test 4

Real words correctly identified	Non-words correctly identified	Real words not identified
<i>Abdomen</i> (abdomen)	<i>Apetitar</i>	<i>Autolimitante</i> (self-limiting)
<i>Blando</i> (soft)	<i>Cannabidiático</i>	<i>Carraspera</i> (itchy throat)
<i>Mascota</i> (pet)	<i>Carrasperante</i>	<i>Gasificar</i> (to gasify)
<i>Cachorro</i> (puppy)	<i>Comalir</i>	<i>Plastilina</i> (modeling clay)
<i>Defecar</i> (to defecate)	<i>Diarreante</i>	<i>Rasurar</i> (to shave)
<i>Diarrea</i> (diarrhea)	<i>Ecocardiante</i>	
<i>Orinar</i> (to urinate)	<i>Gastroenterizar</i>	
<i>Pulga</i> (flea)	<i>Intratoracoide</i>	
<i>Tos</i> (cough)	<i>Mascotar</i>	
<i>Vacuna</i> (vaccine)	<i>Medularer</i>	
	<i>Neurólogoar</i>	
	<i>Quíntuplar</i>	
	<i>Riñonar</i>	
	<i>Tumórico</i>	

4.3.1 Assessing Performance Consistency and Retention Over Time

This section provides the results of the statistical analysis of participants' scores over time. The findings derived from the implementation of the repeated measures ANOVA highlighted that there was no statistical significance in test scores across the four times it was

administered, ($F(3,39) = 1.8, p = 0.163$). Given the absence of a significant effect over time, this also indicates that there was no significant difference observed between explicit and implicit instructional approaches, indicating that neither method produced a statistically distinct impact on participants' vocabulary acquisition.

4.4 Discussion

This study explored how a VR/AI platform impacts the learning and retention of receptive vocabulary in students enrolled in a Spanish for Veterinary Medicine course. To do so, three vocabulary tests were given at three different points during one semester, and a final test was given three months after the end of the semester. With the results given, this discussion is focused on the impact of VR/AI on vocabulary acquisition and retention. Another topic that will be discussed is the differences between the four tests as well as the words and non-words that were recognized. This section concludes with the implications for educational practice.

4.4.1 Impact of VR and AI on Vocabulary Acquisition

While previous research has found that immersive VR may serve as an effective medium for L2 vocabulary development (Legault et al., 2019) and can foster student engagement (Radianti et al., 2020), results of this study showed no statistical significance between test scores across the four time points, which could be explained by different factors. One of these factors could be that the VR platform used in this study was non-immersive. The absence of full immersion could have limited learners' ability to form deeper contextual connections between lexical items and their real-world applications. In addition to the small sample size, the first three tests were conducted over a total of six weeks, and as previous research shows, it is difficult to measure significant proficiency development in such a short period of time (Milton, 2009; Saito et al., 2019). As the results showed no statistically significant differences in test scores across the

four assessments, this suggests that neither explicit nor implicit instructional approaches led to measurable improvements within this timeframe. However, as Schmitt & Schmitt (2020) argued, a combination of explicit instruction and incidental learning opportunities can lead to more robust vocabulary acquisition and retention. As such, integrating VR and AI with conventional vocabulary exercises may offer a promising approach to creating a balanced learning environment, but additional research is needed. Furthermore, the observed variability in word recognition patterns highlights the complexity of vocabulary acquisition, with some words showing prolonged recognition while others exhibited a decline over time. This finding suggests that vocabulary acquisition does not follow a strictly linear progression, as different lexical items may require varying levels of exposure and reinforcement to ensure lasting retention Nation (2022).

The retention assessment administered three months after the final instructional session revealed a slight decrease in scores, with an average score of 26.86 in Test 4 compared to 27.87 in Test 3. However, this difference was not significant, indicating that students' vocabulary knowledge remained stable even three months after the course ended. While VR may contribute to engagement and contextualized learning, its impact on long-term vocabulary retention remains uncertain, which future research should explore.

Changes in how words were classified (i.e., real vs. non-real) were observed across assessments. Four newly recognized real words (*hepático*/hepatic, *abdomen*/abdomen, *esterilizar*/to sterilize, *blando*/soft) emerged from Test 1 to Test 2, while two previously acknowledged terms (*mascota*/pet and *riñón*/kidney) were no longer unanimously classified as real words. Non-word identification remained variable, with seven items designated as non-

words, three of which had been identified in the initial evaluation. Misclassification persisted, as *carraspera* (itchy throat) was again incorrectly labeled as a non-word.

In the third assessment, the overall recognition of real words declined slightly. Among the ten lexical items classified as real, one new term (*gastroenteritis*) was identified, while three previously recognized words (*hepático*/hepatic, *blando*/soft, *riñón*/kidney) were no longer unanimously acknowledged. Non-word classification totaled eleven items, with five newly introduced and three real words misclassified (*carraspera*/itchy throat, *plastilina*/modeling clay, *gasificar*/to gasify), reinforcing inconsistencies in lexical identification. In the final test, all ten identified real words appeared in previous assessments. Four terms (*gastroenteritis*/gastroenteritis, *esterilizar*/to sterilize, *hepático*/hepatic, *riñón*/kidney) lost unanimous classification, showing continued variability in recognition patterns. Non-word identification rose to fifteen, with eight newly introduced lexical items. Misclassification increased, as five real words were incorrectly labeled, including *carraspera* (itchy throat), *plastilina* (modeling clay), *gasificar* (to gasify), *rasurar* (to shave), and *autolimitante* (self-limiting). Despite being able to see these changes across the four tests, it cannot be said that there were significant differences between them.

As evidenced, how participants recognized words changed over time, with some new words being identified correctly while others were mistakenly classified or forgotten. This finding suggests that learners need repeated exposure and reinforcement to solidify their vocabulary (Zarfsaz & Yeganehpour, 2021). As Nation (2022) discussed, vocabulary learning is not linear as some words may be remembered quickly, while others require repeated interaction before they are fully retained. Additionally, the inconsistencies in non-word identification

suggest that participants had difficulties deciding which words were real, questioning whether VR/AI learning environments can support long-term vocabulary retention.

4.4.2 Implication for Educational Practice

The results of this investigation have several implications for pedagogical practice. While prior research has explored VR as a tool for language learning engagement (Kaplan-Rakowski & Wojdyski, 2018; Mayer, 2014), these findings suggest that its integration did not lead to observable improvements in vocabulary retention during the study period. As Schmitt & Schmitt (2020) emphasized, systematic review and vocabulary practice remain essential for retention, regardless of instructional methods. Educators should develop strategies for structured vocabulary review to ensure that learners have opportunities to reinforce lexical knowledge over time. Additionally, this study contributes to a bigger conversation on optimizing language instruction by highlighting the importance of evaluating emerging technologies within pedagogical frameworks.

Emerging technologies, while valuable, may not be sufficient on their own for fostering vocabulary development in LSP. To maximize their impact, educators should adopt hybrid instructional models that blend VR-driven situational learning with systematic vocabulary review grounded in established linguistic theories (Özgün & Sadik, 2023). Hybrid instructional models should account for not just the integration of VR, but also their level of immersion, which may impact engagement and retention. Research indicates that immersive environments enhance engagement, but meaningful learning outcomes depend on structured pedagogical frameworks that support retention and application (Barráez-Herrera, 2022; Lin & Lan, 2015).

4.4.3 Limitations and Future Research

A key limitation of this study was the small sample size, which impacts both the power and reliability of statistical analyses. With fewer participants, detecting meaningful differences across tests becomes increasingly difficult, particularly in longitudinal studies where individual variability can hide broader trends. This limitation restricts the generalizability of the findings, as patterns observed within a small group may not accurately reflect results in a larger population. Additionally, the absence of a control group prevents direct comparisons between instructional methods, making it difficult to determine whether the use of a VR/AI platform or activity types influenced the results. Without this control group, any observed changes could be attributed to external factors and not to the instructional approach. Finally, the intervention was conducted over six weeks, which may be insufficient to capture substantial changes in vocabulary acquisition. Prior research has emphasized that long-term exposure and reinforcement are critical for measurable gains in language proficiency (Schmitt & Schmitt, 2020). Future studies should extend the duration of instruction to allow for more robust conclusions regarding the effects of VR-based vocabulary learning.

CHAPTER 5: DISCUSSION AND CONCLUSION

This research focused on the development and validation of the LexVet-Esp, a vocabulary assessment tool designed to measure receptive vocabulary knowledge amongst students engaged in a Spanish curriculum specifically for veterinary contexts. This assessment was also used to explore the influence of VR and AI on language acquisition and to determine whether explicit vocabulary exercises result in better learning outcomes compared to implicit vocabulary acquisition methods. The results of the different tests were compared before and after the introduction of the VR and AI-based educational activities, as well as contrasting results from instructional weeks focused on explicit vocabulary exercises with those emphasizing implicit acquisition strategies. This approach allowed an investigation of how the educational environment and teaching strategies might impact vocabulary acquisition.

Initial comparison showed some variation in scores, but no statistically significant differences were observed across the four tests. This means that no clear conclusions could be inferred about the effectiveness of explicit versus implicit vocabulary activities or the impact of using a VR/AI platform on vocabulary retention. The small sample size ($n=15$) narrowed the analysis, reducing the ability to detect potential trends. A larger, more diverse participant group could provide a clearer understanding of how instructional methods influence vocabulary development. Moreover, the duration of the study was over six weeks, and a longer time period may be necessary to detect any potential long-term impact of different learning interventions on vocabulary retention. Finally, the absence of a true control group prevents direct comparisons between traditional and technology-enhanced instruction, making it difficult to isolate the effect of the use of VR and AI.

5.1 Validation of LexVet-Esp

The LexVet-Esp was designed to address the specific needs for field-specific vocabulary assessment in Spanish for Veterinary Medicine. With the growing importance of Spanish in veterinary fields, especially due to demographic changes in the U.S. (Brown, 2023; US Census Bureau, 2023, 2024a), this tool offers a more precise evaluation of vocabulary knowledge compared to general tests like the Lextale-Esp (Izura et al., 2014), which do not take into consideration specialized terminology (Sánchez-López et al., 2017). The LexVet-Esp's development involved corpus-based item selection, and the corpus was based on an LNA, which ensured that the vocabulary found in the LexVet-Esp aligned with authentic usage in veterinary practice (El Arbaoui, 2024; Nepravishta & Roseni, 2014). As there was just one data collection site, this limits the generalization of the findings. This consideration is especially relevant when reflecting on previous assessments, such as LexITA and Lextale-Esp, which featured participants from multiple universities and diverse linguistic backgrounds, thus strengthening the validity of their results (Amenta et al., 2020; Izura et al., 2014).

5.2 Considerations for VR, AI, and Vocabulary Retention

VR represents a significant technological innovation in L2 education as it offers language learners the opportunity to practice another language in authentic, real-life scenarios, which enhances engagement and motivation (Chen et al., 2022; Özgün & Sadik, 2023). These types of scenarios can also align with TBLT principles by emphasizing tasks that require the use of specialized vocabulary in meaningful ways (de Oliveira & Lemos, 2022; Naudi, 2023). In conjunction with VR, AI is recognized as a powerful tool in language education. AI technologies can support personalized learning by offering real-time feedback, adaptive content delivery, and context-aware language practice, which has the potential to improve the learner autonomy and

their vocabulary acquisition (Macinska & Vinkler, 2021; Son et al., 2023; Zhu & Wang, 2025). However, the findings of this study do not provide evidence that VR contributed to measurable vocabulary gains. The integration of VR within an L2 classroom has been associated with improved engagement and more contextualized learning experiences (Kaplan-Rakowski & Wojdyski, 2018), and research has emphasized that meaningful exposure to vocabulary in context can support retention (Swales, 2000). It is also relevant to mention that the level of immersion may play a crucial role in determining its effectiveness for vocabulary retention. However, due to this study's timeframe of six weeks, which may have been too brief, and the low number of participants, it was difficult to assess long-term effects adequately. Also, the VR platform used in this study was non-immersive, which may have limited learners' ability to establish meaningful contextual connections short and long-term. Further research incorporating longitudinal assessment methods is necessary to determine whether non-immersive VR-based learning strategies yield measurable improvements.

5.3 Future Research

Future research should expand the applicability of the LexVet-Esp by investigating its relationship with other linguistic competencies, including reading comprehension, listening comprehension, grammatical competence, and speaking fluency. It would also be valuable to compare the LexVet-Esp to other proficiency tests, such as the Lextale-Esp or the *Diploma de Español como Lengua Extranjera* (DELE), to examine convergent validity. Although, it is important to note that these standardized assessments are focused on general Spanish language proficiency and do not emphasize field-specific vocabulary in professional contexts.

Through the examination of these relationships, more insights could be developed related to how receptive vocabulary knowledge interacts with different linguistic competencies,

specifically in professional settings. Findings from such studies may help guide curriculum design and assessment while also identifying vocabulary categories essential for professional competency. Another area for research involves monitoring the progression of test results over the length of the Spanish for Veterinary Medicine certificate program, which consists of four semesters and nine credits, which would measure individual participants' growth in receptive vocabulary over a longer period.

Ultimately, while the present study does not confirm measurable impacts of VR and AI on vocabulary acquisition and retention, future research should continue to explore technology's potential role in education longitudinally. Studies should examine how continued exposure to VR environments, combined with AI-driven feedback and adaptive vocabulary tasks, influences long-term retention and transfer of field-specific vocabulary. Additionally, future research should compare different levels of VR immersion to determine which type of VR environment is better suited for vocabulary acquisition and retention.

Future research efforts should also investigate how different technologies might be adapted to other languages and professional contexts. For instance, assessment tools such as the LexVet-Esp, originally designed for veterinary medicine, could be adapted for other fields, including medicine, engineering, or business, in Spanish or additional languages. With the combination of these new assessment tools with technological advancements, researchers should further examine new approaches to language learning and more effectively evaluate the vocabulary knowledge learners gain in LSP courses.

REFERENCES

- Aghapour, M., & Bockstahler, B. (2022). State of the art and future prospects of virtual and augmented reality in veterinary medicine: A systematic review. *Animals*, 12(24), 3517. <https://doi.org/10.3390/ani12243517>
- Agurto-Cabrera, J. C., & Guevara-Vizcaíno, C. F. (2023). Realidad virtual para la mejora del rendimiento académico en estudiantes de educación superior. *Revista Metropolitana de Ciencias Aplicadas*, 6(2, Suppl.), 233–234.
- Amenta, S., Badan, L., & Brysbaert, M. (2020). LexITA: A quick and reliable assessment tool for Italian L2 receptive vocabulary size. *Applied Linguistics*, 42(2, April 2021), 292–314. <https://doi.org/10.1093/applin/amaa020>
- Anaconda Ortiz, J. D., Millán Rojas, E. E., & Gómez Cano, C. A. (2019). Aplicación de los metaversos y la realidad virtual en la enseñanza. *Entre Ciencia e Ingeniería*, 13(25), 59–67. <https://doi.org/10.31908/19098367.4015>
- Appleby, R. B., & Basran, P. S. (2022). Artificial intelligence in veterinary medicine. *Journal of the American Veterinary Medical Association*, 260(8), 819–824. <https://doi.org/10.2460/javma.22.03.0093>
- Barráez-Herrera, D. P. (2022). Metaversos en el contexto de la educación virtual. *Revista Tecnológica-Educativa Docentes 2.0*, 13(1), 11–19. <https://doi.org/10.37843/rted.v13i1.300>
- Barsalou, L. W. (2008). Grounded cognition. *Annual Review of Psychology*, 59, 617–645. <https://doi.org/10.1146/annurev.psych.59.103006.093639>

- Barsalou, L. W. (2023). Implications of grounded cognition for conceptual processing across cultures. *Topics in Cognitive Science*, 15(4), 648–656. <https://doi.org/10.1111/tops.12661>
- Basturkmen, H. (2010). Analysing needs. In H. Basturkmen, *Developing Courses in English for Specific Purposes* (pp. 17–35). Palgrave Macmillan UK.
https://doi.org/10.1057/9780230290518_2
- Blyth, C. (2018). Immersive technologies and language learning. *Foreign Language Annals*, 51(1), 225–232. <https://doi.org/10.1111/flan.12327>
- Borawski, G. (2019). Optimizing L2 vocabulary acquisition: Applied linguistic research. Paper presented at the *Purdue Linguistics, Literature, and Second Language Studies Conference*, Purdue University, West Lafayette, IN.
<https://docs.lib.purdue.edu/plcc/2019/papers/1>
- Branden, K. V. D. (Ed.). (2006). *Task-Based language education: From theory to practice*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511667282>
- Brown, A. (2023, July 7). About half of U.S. pet owners say their pets are as much a part of their family as a human member. *Pew Research Center*. <https://www.pewresearch.org/short-reads/2023/07/07/about-half-us-of-pet-owners-say-their-pets-are-as-much-a-part-of-their-family-as-a-human-member/>
- Cédric, S. (2021). Technology-Mediated ESP learning and teaching (TESPLAT): From language learning theory to ICT integration principles. In A. Turula & A. Buczek-Zawila (Eds.), *CALL for Background*. Peter Lang. <https://doi.org/10.3726/b18284>

- Chacón Medina, A. (2007). La tecnología educativa en el marco de la didáctica. In J. A. Ortega Carrillo & A. Chacón Medina (Eds.), *Nuevas tecnologías para la educación en la era digital* (pp. 25–41). Pirámide.
- Chen, B., Wang, Y., & Wang, L. (2022). The effects of virtual reality-assisted language learning: A meta-analysis. *Sustainability*, *14*(6), 3147. <https://doi.org/10.3390/su14063147>
- Czerkawski, B., & Berti, M. (2020). Language learning in the 21st century: Current status and future directions. In B. Dupuy & M. Grosbois (Eds.), *Language learning and professionalization in higher education: Pathways to preparing learners and teachers in/for the 21st century* (1st ed., pp. 11–35). Research-publishing.net. <https://doi.org/10.14705/rpnet.2020.44.1100>
- de Oliveira, M., & Lemos, S. (2022). A task-based needs analysis of university-level portuguese for business. *Hispania*, *105*(3), 401–422.
- Dodigovic, M. (1993). Interdisciplinary: Computer assisted linguistic research and the development of LSP courseware. *CALICO Journal*, *10*(4), 5–16.
- El Arbaoui, F. Z. (2024). Implementing task-based approach in ESP education: Business schools as a case study. *Acuity: Journal of English Language Pedagogy, Literature, and Culture*, *9*(2), 252–266. <https://doi.org/10.35974/acuity.v9i2.3178>
- Ellis, R. (2003). *Task-based language learning and teaching*. Oxford University Press.
- Fitzpatrick, T. (2007). Productive vocabulary tests and the search for concurrent validity. In H. Daller, J. Milton, & J. Treffers-Daller (Eds.), *Modelling and Assessing Vocabulary*

Knowledge (1st ed., pp. 116–132). Cambridge University Press.

<https://doi.org/10.1017/CBO9780511667268.009>

Forehand, E. E., Zeller, S. K., & Frey, D. M. (2023). Spanish for veterinarians part 2: A survey gauging student perspectives on a profession-specific language learning experience.

Journal of Veterinary Medical Education, 50(5), 542–553. <https://doi.org/10.3138/jvme-2022-0137>

Georgy, C. M. (2023). Implementing task-based language teaching in an ESP classroom: Design and implications. *Journal of English Language and Literature*, 10(02), 116–123.

<https://doi.org/10.54513/JOELL.2023.10212>

Godwin-Jones, R. (2011). Emerging technologies: Mobile apps for language learning.

Language Learning & Technology, 15(2), 2–11.

Gómez de Enterría, J. (2009). *El español lengua de especialidad: Enseñanza y aprendizaje*. Arco Libros.

Ha, H. T. (2021). Exploring the relationships between various dimensions of receptive vocabulary knowledge and L2 listening and reading comprehension. *Language Testing in Asia*, 11(1), 20.

<https://doi.org/10.1186/s40468-021-00131-8>

Hattani, H. A. (2020). The implementation of task-based approach in ESP instruction: Teachers and students' perceptions. *International Arab Journal of English for Specific Purposes*,

3(1), 67-82. <https://doi.org/10.34874/PRSM.iajesp-vol3iss1.23539>

Helms, L., Pitner, H., & Adams, R. (2023). Spanish for nursing: A needs analysis for a Spanish for nursing course. *Revista de Lenguas Para Fines Específicos*, 29, 74–93.

- Holden, C. L., & Sykes, J. M. (2011). Leveraging mobile games for place-based language learning. *International Journal of Game-Based Learning*, 1(2), 1–18.
<https://doi.org/10.4018/ijgbl.2011040101>
- Ibáñez, M.-B., & Delgado-Kloos, C. (2018). Augmented reality for STEM learning: A systematic review. *Computers & Education*, 123, 109–123.
<https://doi.org/10.1016/j.compedu.2018.05.002>
- Izura, C., Cuetos, F., & Brysbaert, M. (2014). Lextale-Esp: A test to rapidly and efficiently assess the Spanish vocabulary size. *Psicológica*, 35(1), 49–66.
<https://doi.org/10.1037/t47086-000>
- Kaplan-Rakowski, R., & Wojdyski, T. (2018). Students' attitudes toward high-immersion virtual reality assisted language learning. In *Future-proof CALL: language learning as exploration and encounters – short papers from EUROCALL 2018*. Social Science Research Network. <https://papers.ssrn.com/abstract=3255611>
- Karpicke, J. D., & Roediger III, H. L. (2008). The critical importance of retrieval for learning. *Science*, 319(5865), 966–968. <https://doi.org/10.1126/science.1152408>
- Khatib, M., & Dehghankar, A. (2018). The impact of task-based language teaching on ESP learners' productive skills: From task-based instruction to investigation of learners' and instructors' attitudes toward the course. *Issues in Language Teaching*, 7(2), 1–28.
<https://doi.org/10.22054/ilt.2019.46324.422>
- Klimova, B. (2021). Use of virtual reality in non-native language learning and teaching. *Procedia Computer Science*, 192, 1385–1392.
<https://doi.org/10.1016/j.procs.2021.08.141>

- Kukulska-Hulme, A., & Shield, L. (2008). An overview of mobile assisted language learning: From content delivery to supported collaboration and interaction. *ReCALL*, 20(3), 271–289. <https://doi.org/10.1017/S0958344008000335>
- Lafford, B. (2012). Languages for Specific Purposes in the United States in a Global Context: Commentary on Grosse and Voght (1991) Revisited. *The Modern Language Journal*, 96(s1), 1–27. <https://doi.org/10.1111/j.1540-4781.2012.01294.x>
- Lafford, B. (2022). *Spanish for specific purposes in the United States in a global contest: Moving the field forward through communication, connections and collaboration*. LSP Research Cluster Roundtable, Southern Methodist University.
- Lafford, B. (2024). Spanish for specific purposes in the United States: Trends, challenges, solutions, and future directions. In M. Gradoville & S. McKinnon (Eds.), *Issues in Hispanic and Lusophone Linguistics* (Vol. 41, pp. 268–288). John Benjamins Publishing Company. <https://doi.org/10.1075/ihll.41.11laf>
- Landau, R. E. (2013). *Assessing the preparedness of the veterinary profession to communicate with limited English proficient Spanish-speaking pet owners* [Doctoral dissertation, Purdue University]. ProQuest Dissertations & Theses. <https://www.proquest.com/docview/1511024575/abstract/ECE3B0330AC498EPQ/1>
- Landau, R. E., Beck, A., Glickman, L. T., Litster, A., Widmar, N. J. O., & Moore, G. E. (2015). Survey of US veterinary students on communicating with limited English proficient Spanish-speaking pet owners. *Journal of Veterinary Medical Education*, 42(4), 324–331. <https://doi.org/10.3138/jvme.0215-025R>

- Larkin, M. (2024). *Pet population continues to increase while pet spending declines*. American Veterinary Medical Association. <https://www.avma.org/news/pet-population-continues-increase-while-pet-spending-declines>
- Laufer, B., & Nation, P. (1999). A vocabulary-size test of controlled productive ability. *Language Testing*, 16(1), 33–51. <https://doi.org/10.1177/026553229901600103>
- Legault, J., Zhao, J., Chi, Y.-A., Chen, W., Klippel, A., & Li, P. (2019). Immersive virtual reality as an effective tool for second language vocabulary learning. *Languages*, 4(1), Article 1. <https://doi.org/10.3390/languages4010013>
- Lemhöfer, K., & Broersma, M. (2012). Introducing LexTALE: A quick and valid lexical test for advanced learners of English. *Behavior Research Methods*, 44(2), 325–343. <https://doi.org/10.3758/s13428-011-0146-0>
- Li, Y., Ying, S., Chen, Q., & Guan, J. (2022). An experiential learning-based virtual reality approach to foster students' vocabulary acquisition and learning engagement in English for Geography. *Sustainability*, 14(22), 15359. <https://doi.org/10.3390/su142215359>
- Lin, T.-J., & Lan, Y.-J. (2015). Language learning in virtual reality environments: Past, present, and future. *Journal of Educational Technology & Society*, 18(4), 486–497.
- Long, M. H. (1985). A role for instruction in second language acquisition: Task-based language teaching. In K. Hyltenstam & M. Pienemann (Eds.), *Modeling and assessing second language development* (pp. 77–99). Multilingual Matters.
- Long, M. H. (Ed.). (2005). *Second language needs analysis*. Cambridge University Press.

- Macinska, S., & Vinkler, A. (2021). *The role of AI in language learning: Research evidence and strategies for success*. Cambridge Papers in English Language Education.
https://www.cambridge.org/sites/default/files/media/documents/The_Role_of_AI_in_Language_Learning_Research_Evidence_and_Strategies.pdf
- Malicka, A., Gilabert Guerrero, R., & Norris, J. M. (2019). From needs analysis to task design: Insights from an English for specific purposes context. *Language Teaching Research*, 23(1), 78–106. <https://doi.org/10.1177/1362168817714278>
- Masrai, A. (2023). Lexical knowledge and L2 general language proficiency: Collocational competence and vocabulary size as determinants of lexical knowledge. *Cognitive Processing*, 24(2), 289–300. <https://doi.org/10.1007/s10339-022-01120-2>
- Mayer, R. E. (2009). *Multimedia Learning* (2nd ed.). Cambridge University Press.
<https://doi.org/10.1017/CBO9780511811678>
- Mayer, R. E. (2014). Cognitive theory of multimedia learning. In R. E. Mayer (Ed.), *The Cambridge Handbook of Multimedia Learning* (2nd ed., pp. 43–71). Cambridge University Press. <https://doi.org/10.1017/CBO9781139547369.005>
- Mcbride-Chang, C., Tardif, T., Cho, J.-R., Shu, H., Fletcher, P., Stokes, S. F., Wong, A., & Leung, K. (2008). What’s in a word? Morphological awareness and vocabulary knowledge in three languages. *Applied Psycholinguistics*, 29(3), 437–462.
<https://doi.org/10.1017/S014271640808020X>
- Miller de Rutté, A. (2024a). The design and implementation of virtual reality simulations in Spanish for the Health Professions: What are learners’ perceptions? *Computer Assisted Language Learning*, 1–21. <https://doi.org/10.1080/09588221.2024.2393325>

- Miller de Rutté, A. (2024b). Virtual reality, medical Spanish, and the L2 motivational self system: A case study. *Technology, Knowledge and Learning*, 29, 827–841.
<https://doi.org/10.1007/s10758-023-09710-5>
- Miller De Rutté, A. M., Kentengian, I. M., Pérez, M. E., & Ortega, P. (2024). The current state of undergraduate medical Spanish: Results of a national survey. *Foreign Language Annals*, 57(1), 27–47. <https://doi.org/10.1111/flan.12731>
- Milton, J. (2009). *Measuring second language vocabulary acquisition*. Multilingual Matters.
<https://doi.org/10.21832/9781847692092>
- Moreno Martínez, N. M., & Galván Malagón, M. C. (2020). Realidad aumentada y realidad virtual para la creación de escenarios de aprendizaje de la lengua inglesa desde un enfoque comunicativo. *Revista DIM: Didáctica, Innovación y Multimedia*, 38, 2.
- Motlagh, F. A., Jafari, A. S., & Yazdani, Z. (2014). A general overview of task-based language teaching (TBLT), from theory to practice. *International Journal of Language and Linguistics*, 2(5–1), 1–11.
- Nation, I. S. P. (1990). *Teaching and learning vocabulary*. Newbury House Publishers.
- Nation, I. S. P. (2022). *Learning vocabulary in another language* (3rd ed.). Cambridge University Press. <https://doi.org/10.1017/9781009093873>
- Nation, I. S. P., & Beglar, D. (2007). A vocabulary size test. *The Language Teacher*, 37(7), 9–13.
- Naudi, A. A. (2023). A task-based road map to teaching business Spanish with word combinations. *Global Business Languages*, 23, 40–57. <https://doi.org/10.4079/gbl.v23.4>

- Nazari, M. (2020). Language for specific purposes teacher education: A scoping review. *MEXTESOL Journal*, 44(4). <https://eric.ed.gov/?id=EJ1311073>
- Nepravishta, A., & Roseni, E. (2014). The impact of task-based language teaching in ESP education in the Agricultural University of Tirana. *Albanian Journal of Agricultural Sciences*, 13(4), 1–5.
- Nunan, D. (2014). Task-based teaching and learning. In M. Celce-Murcia, D. M. Brinton, & M. A. Snow (Eds.), *Teaching English as a Second or Foreign Language* (4th ed., pp. 455–470). Sherrise Roher.
- Özgün, O., & Sadik, O. (2023). Implementation of VR technologies in language learning settings: A systematic literature review. *Educational Policy Analysis and Strategic Research*, 18(4), 32–61.
- Pachler, N., Bachmair, B., & Cook, J. (2010). *Mobile Learning*. Springer US. <https://doi.org/10.1007/978-1-4419-0585-7>
- Park, J. (2022). Vocabulary learning through a daily task of cooking in the Digital Kitchen. *Language Learning & Technology*, 26(1), 1–22.
- Pastor Cesteros, S. (2013). Enseñanza de español con fines profesionales y académicos y aprendizaje por contenidos en contexto universitario. In B. Montoya Abat & A. Mas i Miralles (Eds.), *Studia linguistica in honorem Francisco Gimeno Menéndez* (pp. 765–786). Universitat d'Alacant. <http://rua.ua.es/dspace/handle/10045/48800>

- Peixoto, B., Pinto, R., Melo, M., Cabral, L., & Bessa, M. (2021). Immersive virtual reality for foreign language education: A PRISMA systematic review. *IEEE Access*, 9, 48952–48962. <https://doi.org/10.1109/ACCESS.2021.3068858>
- Pérez, O. A. (2021). Spanish for STEM in US higher education: A historical review and future perspectives. *Global Business Languages*, 21. <https://doi.org/10.4079/gbl.v21.1>
- Pham, X.-L., Chen, G.-D., Nguyen, T.-H., & Hwang, W.-Y. (2016). Card-based design combined with spaced repetition: A new interface for displaying learning elements and improving active recall. *Computers & Education*, 98, 142–156. <https://doi.org/10.1016/j.compedu.2016.03.014>
- Pignot-Shahov, V. (2012). Measuring L2 receptive and productive vocabulary knowledge. *University of Reading Language Studies Working Papers*, 4, 37–45.
- Qasserras, L. (2023). Systematic review of communicative language teaching (CLT) in language education: A balanced perspective. *European Journal of Education and Pedagogy*, 4(6), 17–23. <https://doi.org/10.24018/ejedu.2023.4.6.763>
- Radianti, J., Majchrzak, T. A., Fromm, J., & Wohlgenannt, I. (2020). A systematic review of immersive virtual reality applications for higher education: Design elements, lessons learned, and research agenda. *Computers & Education*, 147, 103778. <https://doi.org/10.1016/j.compedu.2019.103778>
- Reinders, H., & Hubbard, P. (2013). CALL and autonomy: Affordances and constraints. In M. Thomas, H. Reinders, & M. Warschauer (Eds.), *Contemporary Computer-Assisted Language Learning* (pp. 359–375). Continuum.

- Reinders, H., & Wattana, S. (2015). Affect and willingness to communicate in digital game-based learning. *ReCALL*, 27(1), 38–57. <https://doi.org/10.1017/S0958344014000226>
- Richards, J. C., & Rodgers, T. S. (2001). *Approaches and methods in language teaching* (18. print). Cambridge Univ. Press.
- Robinson, P. (2011). Task - based language learning: A review of issues. *Language Learning*, 61(s1), 1-36. <https://doi.org/10.1111/j.1467-9922.2011.00641.x>
- Roche, T., & Harrington, M. (2013). Recognition vocabulary knowledge as a predictor of academic performance in an English as a foreign language setting. *Language Testing in Asia*, 3(1), 12. <https://doi.org/10.1186/2229-0443-3-12>
- Saito, K., Suzukida, Y., & Sun, H. (2019). Aptitude, experience, and second language pronunciation proficiency development in classroom settings: A longitudinal study. *Studies in Second Language Acquisition*, 41(1), 201–225. <https://doi.org/10.1017/S0272263117000432>
- Salazar, L., Martinez Aguiriano, A. J., Pietrosevoli, S., & Garcia, A. (2024). Developing courses of Spanish for Specific Purposes in agriculture to bridge the communication gap between the hispanic workforce and English-speaking veterinary and animal sciences students. *Animals*, 14(24), 3639. <https://doi.org/10.3390/ani14243639>
- Samaraweera, S. G. S. (2025). The influence of morphological awareness and vocabulary depth on academic English writing skills of undergraduates. *Sri Lanka Journal of Social Sciences and Humanities*, 5(1), 35–45. <https://doi.org/10.4038/sljssh.v5i1.128>

- Sanatullova-Allison, E. (2014). Memory retention in second language acquisition and instruction: Insights from literature and research. *IAFOR Journal of Language Learning*, *1*(1). <https://doi.org/10.22492/ijll.1.1.02>
- Sánchez-López, L., Long, M. K., & Lafford, B. (2017). New directions in LSP research in US higher educations. In M. K. Long (Ed.), *Language for specific purposes: Trends in curriculum development* (pp. 13–34). International Symposium on Languages for Specific Purposes, Washington, DC. Georgetown University Press.
- Schmitt, N., & Meara, P. (1997). Researching vocabulary through a word knowledge framework. *Studies in Second Language Acquisition*, *19*(1), 17–36.
<https://doi.org/10.1017/S0272263197001022>
- Schmitt, N., & Schmitt, D. (2020). *Vocabulary in language teaching | Applied linguistics and second language acquisition* (2nd ed.). Cambridge University Press.
<https://doi.org/10.1017/9781108569057>
- Skehan, P. (1998). *A cognitive approach to language learning*. Oxford University Press.
- Sobkowich, K. E. (2025). Demystifying artificial intelligence for veterinary professionals: Practical applications and future potential. *American Journal of Veterinary Research*, *86*(S1), S6–S15. <https://doi.org/10.2460/ajvr.24.09.0275>
- Son, J.-B., Ružić, N. K., & Philpott, A. (2023). Artificial intelligence technologies and applications for language learning and teaching. *Journal of China Computer-Assisted Language Learning*, *5*(1), 94-112. <https://doi.org/10.1515/jccall-2023-0015>

- Sun, D., Chen, Z., & Zhu, S. (2023). What affects second language vocabulary learning? Evidence from multivariate analysis. *Frontiers in Education*, 8, 1210640.
<https://doi.org/10.3389/feduc.2023.1210640>
- Sun, W. (2023). The impact of automatic speech recognition technology on second language pronunciation and speaking skills of EFL learners: A mixed methods investigation. *Frontiers in Psychology*, 14. <https://doi.org/10.3389/fpsyg.2023.1210187>
- Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. Gass & C. Madden (Eds.), *Input in second language acquisition* (pp. 235–253). Newbury House.
- Swales, J. M. (2000). Languages for Specific Purposes. *Annual Review of Applied Linguistics*, 20, 59–76. <https://doi.org/10.1017/S0267190500200044>
- Sykes, J. M., & Reinhardt, J. (2012). *Language at play: Digital games in second and foreign language teaching and learning*. Pearson.
- Thorne, S. L. (2008). Mediating technologies and second language learning. In J. Coiro, M. Knobel, L. C., & D.J Leu (Eds.), *Handbook of research on new literacies* (pp. 417–449). Lawrence Erlbaum Associates.
- Trace, J., Hudson, T., & Brown, J. D. (Eds.). (2015). *Developing courses in languages for specific purposes*. National Foreign Language Resource Center.
- US Census Bureau. (2023). *U.S. Census Bureau Releases 2018-2022 ACS 5-Year Estimates*.
<https://www.census.gov/programs-surveys/acs/news/updates/2023.html>

- US Census Bureau. (2024a). *American Community Survey (ACS)*. Census.Gov.
<https://www.census.gov/programs-surveys/acs>
- US Census Bureau. (2024b). *How immigrants have dispersed throughout the country*.
Census.Gov. <https://www.census.gov/library/stories/2024/04/where-do-immigrants-live.html>
- US Census Bureau. (2024c). *U.S. Foreign-Born Population*. Census.Gov.
<https://www.census.gov/library/visualizations/interactive/foreign-born-population-2018-2022.html>
- Valero Franco, C., & Berns, A. (2023). Desarrollo de apps de realidad virtual y aumentada para enseñanza de idiomas: Un estudio de caso. *RIED-Revista Iberoamericana de Educación a Distancia*, 27(1), 163–185. <https://doi.org/10.5944/ried.27.1.37668>
- Wetherbee, K. (2023, August 3). *AI in language learning: Today's uses, benefits, and concerns*. Langoly. <https://www.langoly.com/ai-language-learning/>
- Xu, X., Kilroy, D., Kumar, A., Iqbal, M. Z., Mangina, E., & Campbell, A. G. (2023). Mixed reality applications in tertiary veterinary education: a systematic review. In Y. Cai, E. Mangina, & S. L. Goei (Eds.), *Mixed Reality for Education* (pp. 241–264). Springer Nature Singapore. https://doi.org/10.1007/978-981-99-4958-8_10
- Yu, Q., Vincent, A., Merfeld-Langston, A. L., & Pérez-Muñoz, C. (2020). Collaborating across organizational lines in Language for Specific Purposes. *Global Business Languages*, 20, 2-17. <https://doi.org/10.4079/gbl.v20.1>

- Zarfsaz, E., & Yeganehpour, P. (2021). The impact of different context levels on vocabulary learning and retention. *Shanlax International Journal of Education*, 9(4), 24–34.
- Zeller, S., & Sarasty, E. (Forthcoming). Improving access to veterinary care: A language needs analysis and curricular design in Spanish for small animal veterinary care. In *Spanish for Specific Purposes in the U.S.: From language needs analysis to evidence-based curricular design and implementation*. Routledge.
- Zeller, S. K., Frye, M. A., & Frey, D. M. (2023). Spanish for veterinarians part 1: An approach to weaving Spanish language education into DVM curricula. *Journal of Veterinary Medical Education*, 50(5), 515–525. <https://doi.org/10.3138/jvme-2022-0136>
- Zeller, S., & Velázquez-Castillo, M. (2018). Spanish for animal health and care: A case for expanding the LSP curriculum. In C. King Ramírez & B. Lafford (Eds.), *Transferable skills for the 21st century: Preparing students for the workplace through world languages for specific purposes* (pp. 287–316). Sabio Books LLC.
- Zeller, S., Velazquez-Castillo, M., & Roman-Muniz, I. N. (2016). Spanish for animal health and care: Toward a certificate program in field-specific Spanish. *Journal of Animal Science*, 94, 854–855.
- Zhu, M., & Wang, C. (2025). A systematic review of research on AI in language education: Current status and future implications. *Language Learning & Technology*, 29(1), 1–29.

APPENDICES

Appendix A: Complete list of 180 terms used to develop the LexVet-Esp

1. Veneno	25. Torácico	49. Ganglionizar	73. Vestibular
2. Lipémico	26. Alérgico	50. Vomitar	74. Microcitar
3. Pulgario	27. Hematocito	51. Ecocardio	75. Antipulga
4. Deshidratar	28. Hígado	52. Fosfatar	76. Orinario
5. Abdominado	29. Desplacer	53. Fosfatasa	77. Neurólogo
6. Motilidad	30. Epiléptico	54. Comer	78. Defecar
7. Tráquear	31. Hez	55. Entubar	79. Medularer
8. Apetito	32. Metabolizar	56. Hospitalizarizar	80. Blandecer
9. Médula	33. Alérgizer	57. Motilante	81. Neurólogar
10. Convaleciente	34. Oler	58. Apetitar	82. Intratorácicode
11. Enronchado	35. Pepas	59. Ladiar	83. Desinflamantear
12. Desplacante	36. Gastroenteritis	60. Encapsular	84. Ganglio
13. Ganglionizar	37. Cremaciónico	61. Cachorro	85. Respirar
14. Pulga	38. Remisión	62. Higidizar	86. Ingestar
15. Abdomen	39. Hezoso	63. Tos	87. Metabolizarar
16. Ingestario	40. Gasificarar	64. Tráquea	88. Vómito
17. Linfocito	41. Vestibulante	65. Botica	89. Creatina
18. Anestesiático	42. Infección	66. Tumor	90. Nasolagrimante
19. Radiografía	43. Autolimitantear	67. Defecario	91. Esterilizazir
20. Parásito	44. Deshidratable	68. Creatinar	92. Diarreante
21. Bronquitis	45. Sonda	69. Pulmón	93. Olientar
22. Gasificar	46. Carraspera	70. Antibióticar	94. Cannabidiol
23. Mirántico	47. Carrasperante	71. Hospitalizar	95. Mirar
24. Cannabidiático	48. Respirario	72. Riñonar	

96. Intratorácica
97. Tórax
98. Hipoalergénicos
99. Tosar
100. Acostado
101. Torácicus
102. Acostante
103. Otico
104. Vacunoso
105. Hepático
106. Paralizarar
107. Sedar
108. Esterilizar
109. Alérgar
110. Radiográfílo
111. Desinflamatorio
112. Perrerar
113. Ladilla
114. Naso-lagrimal
115. Ecocardiante
116. Gánglio
117. Torácix
118. Comalir
119. Parálisis
120. Riñón
121. Pulmonarar

122. Glándula
123. Brinquir
124. Plastilina
125. Mascota
126. Rasurar
127. Hemogramar
128. Plastilinar
129. Sedarizar
130. Quíntuplar
131. Popó
132. Bronquitar
133. Peposo
134. Gastroenterizar
135. Encapsulante
136. Cachorro
137. Hipoalergénicar
138. Convaleciantar
139. Tableta
140. Enfermar
141. Anestesia
142. Parasíticoz
143. Entubarar
144. Rasurarar
145. Orinar
146. Linfocitar

147. Glándulico
148. Quíntuple
149. Diarrea
150. Epilépticar
151. Vacuna
152. Infeccionar
153. Sondearar
154. Hemograma
155. Óticante
156. Vomitor
157. Microcítico
158. Popizar
159. Perrera
160. Vomitablear
161. Antipulgario
162. Lipémosis
163. Tabletear
164. Autolimitante
165. Hematocitar
166. Enfermarar
167. Antibiótico
168. Brinco
169. Mascotar
170. Glóbulante
171. Venener
172. Boticar
173. Irritarar
174. Glóbulo

175. Tumórico
176. Enroncharar
177. Remisar
178. Blando
179. Irritar
180. Cremación

Appendix B: Part 1, Instructions, and Example Included in the LexVet-Esp

1. Name
 2. What Spanish class(es) are you currently enrolled in?
 3. What is your race/ethnicity?
 4. What is your age?
 5. What is your primary language? / What language did you speak at home growing up?
- The test includes 90 sequences of letters which may be existing Spanish words or not;
 - Select the words you know (even if you would not be able recall their exact meaning) by ticking the corresponding box;
 - Be careful: errors are penalized! Trying to increase your score by selecting sequences you've never seen before can worsen your final score!
 - For example:

Depiste	
Sí	✓
Cone	
Calpar	
Joten	
Sacapuntas	✓
Priba	
Pelasula	
Bien	✓
Casa	✓
Lejo	
Pretantas	