A Systematic Review of Virtual Reality in the Acquisition of Second Language

https://doi.org/10.3991/ijet.v17i15.31781

Francisco de Paula Forero Pataquiva[1], Blanka Klimova[2](✉)
[1] State University of Campinas, Barão Geraldo, Brazil
[2] University of Hradec Kralove, Hradec Kralove, Czechia
blanka.klimova@uhk.cz

Abstract—Innovations in the fields of virtual reality (VR) and augmented reality (AR) and the availability and usability improvements seen in recent years in these technologies have opened new possibilities in the field of foreign language learning. Using the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) framework, this study conducted a review of seven recent original articles with empirical data focused on the second language (L2) acquisition using immersive reality technologies. The main aim of this study was to analyze what language skills and what proficiency of students' language level benefited most from the use of these new technologies. In addition, the authors explored what kind of VR or AR was used since the term may refer to a series of technologies that do not mean the same. For instance, VR might mean using headset displays, or it may refer to the use of interactive environments. The results indicate that the current VR and AR tools show a positive impact on L2 acquisition when compared to traditional learning methods, such as book-based learning. In addition, the findings reveal that VR and AR can be used among students with zero previous knowledge up to an intermediate and advanced level of fluency, with any L2 language. Nevertheless, only listening and reading skills play a role in the acquisition of L2 using virtual reality. However, as with any recent research field, it should be noted that there are several methodological limitations and that it is necessary to consider the achieved results critically.

Keywords—second language acquisition, language learning, applied linguistics, virtual reality

1 Introduction

New technologies are of particular interest in the field of linguistics, mainly applied linguistics [1]. The reason is that they have been shown to have a positive effect on students' achievement results, increasing their receptiveness and attitude toward learning in general, possibly due to the novelty effect or the actual advantages of new technologies [2][3]. Because of advancements in technology and the increase in hardware capable of running these applications (apps), virtual reality and augmented reality have gained prominence as feasible solutions for studying foreign languages in recent years (mainly smartphones [4]).

With the widespread adoption of smartphones, there has been a tendency of "gamification" in the language learning sector, which leverages the resources provided by technology to offer a better experience for the learner [5][6]. As a result, the employment of VR and AR might bridge the gap between conventional approaches that lack the immersion solely available through the technologies described above [7][8]. Given the benefits that new tools can provide, this systematic review aims to examine the original experimental studies and their findings on the potential of VR/AR as a tool for second language learning. Furthermore, the authors explore the implications for foreign language teaching and learning and the existing limitations of the detected studies.

As with any field with a short lifespan, the results of any of the articles examined in this review should not be regarded as an absolute conclusion but rather as a snapshot of the current state of the technology and a starting point of what is to come in the field of language acquisition and immersive technologies, such as VR and AR. The need for such a study arose from the observation that, due to the developing nature of this topic, no reviews of this type involving the junction of VR/AR and L2 acquisition had been written. Therefore, the key research questions are as follows:

1. What is the definition of VR and AR technology?
2. Which language skills can benefit the most from VR and/or AR technology, and at what level of students' language proficiency?
3. What are the implications for foreign language teaching and learning?

This article is divided into five main sections, the first of which is this introduction. The following part, the methodology section, explains the selection procedure following the PRISMA framework. After that, in the result section, the information of the detected articles for the final review is thoroughly described. In the fourth section, the discussion is provided utilizing the collected material to address concerns about the potential consequences of VR/AR on the learning experience. Finally, the conclusion provides a summary of the essential findings and the field's future.

2 Methodology

This systematic review was prepared using the Preferred Reporting Items for Systematic Reviews and Meta-Analyses approach. For filtering the articles included in the final evaluation, PRISMA involves four stages: identification, screening, eligibility, and inclusion.

After establishing the methodology to search for the literature, it was essential to define the primary questions that this study would address. Given the theme of language acquisition, it was of particular interest to determine which language skills (listening, reading, speaking, and writing) benefited the most from immersion technology and at what degree of proficiency.

Furthermore, it was also crucial to clarify the type of VR/AR because there is no standardized form for this technology. Some rely only on a head-mounted display, and others incorporate interactive aspects through a controller.

For the initial phase, queries were conducted in SCOPUS and the Web of Science with keywords related to virtual reality, augmented reality, and language learning, including synonyms for these keywords. The query was conducted using pairs of words related to VR/AR and a synonym for language learning. Later the operators "OR" and "AND" were used to group all the combinations in a single search in each database.

For the first step, identification, having obtained the search results from the platforms, it was necessary to filter which results were worthwhile. For this purpose, in the screening process, the accepted articles were excluded if they did not meet the criteria of this review. The selection criteria were established, considering the topic of the current investigation to get an array of research papers that could help answer the main points of this systematic review.

For the research papers to be accepted, they had to have empirical data results exploring the role of VR/AR on L2 acquisition. The inclusion criteria are explained as follows:

- Only Peer-reviewed journal articles were included to guarantee the quality of the research.
- Only articles with experimental results were included.
- Only articles published in English were considered since this is the common language of the authors of this article.

In the eligibility process, each article was evaluated to check if it met the characteristics demanded by such research. If this was the case, the study was included in this review. To better describe the steps used in the PRISMA methodology and visualize how it was applied for this review, please consult Figure 1 below.

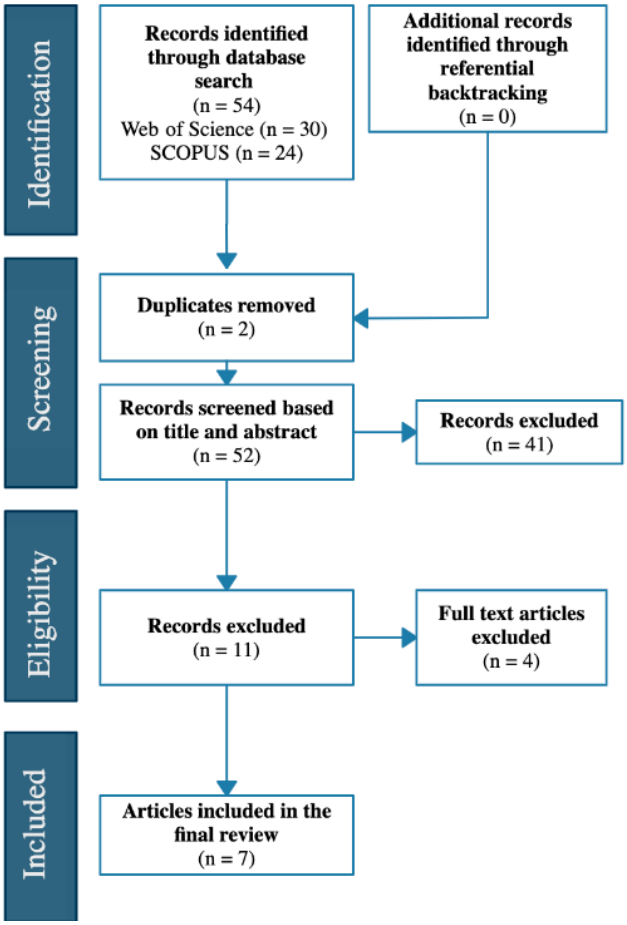

Fig. 1. PRISMA flow chart

3 Results

3.1 Initial classification

After conducting all the refinement steps of the PRISMA framework, seven studies were left. First of all, it is essential to describe the languages studied, the proficiency level, and the type of immersion technology used. Concerning the target languages, the studies focused on Italian, Basque, Japanese, and English. As for the origin of publications, two articles were published in England, two in the United States, two in Switzerland, and one in Egypt. This data shows a moderate level of diversity of languages studied, even if most experiments have English as a second language and are published in countries with English as an official language.

Concerning the educational level of participants involved in the studies, out of the seven articles, four were only made exclusively for undergraduates, two were mixed groups with non-students or students from vocational schools and undergraduates, and finally, just one involved eleventh-graders.

For the initial language proficiency in each study, not all the metrics reported had an equivalency to the Common European Framework of Reference for Languages, which is considered a standard. For both the Italian and Basque languages, there was reported "little or no knowledge of the language." In the case of the Japanese language study, the findings show that participants were mainly at levels N1-N2 and a few at N4 or below (N1 corresponding to the C1 CEFR level, N2 to B2, N3 to B1, N4 to A2 approximately). For the English studies, the findings reveal that two studies included participants with an intermediate level as taught in universities without specifying a CEFR level. Two described their participants as having B1 and B2 proficiency, respectively. Table 1 provides a more detailed overview of these findings.

Table 1. Initial classification

Title	Authors name	The educational level of participants	Target Language	Proficiency level
[1]	I. Nicolaidou, P. Pissas and D. Boglou	Undergraduate	Italian	Little or no knowledge
[2]	C. Repetto, A. Di Natale, D. Villani, S. Triberti, S. Germagnoli and G. Riva	Eleventh-graders	English	B1
[3]	A. Ibrahim, B. Huynh, J. Downey, T. Hollerer, D. Chun and J. O'donovan	Mix of undergraduate and non-students.	Basque	Little or no knowledge
[4]	L. Ma	Undergraduate	English	Intermediate
[5]	R. Shadiev, J. Yu and W. Sintawati	Mix of undergraduate and vocational.	English	Intermediate
[6]	M. Urun, H. Aksoy and R. Comez	Undergraduate	English	B2
[7]	C. Wang, J. Hung and H. Chen	Undergraduate	Japanese	N1-N4

3.2 Technologies used for the experimental groups

As far as the description of the VR setups used in the experiments is concerned, there was no uniform set of technologies. But they could be grouped into two main categories: interactive game environments with a head-mounted VR display and 360° videos recorded from the first-person perspective and shown through a display.

Of the eligible articles, three were not interactive environments, having the use of the headset as its main innovation, and the other four incorporated special interactive environments where subjects could interact by walking or moving their perspective to different settings, in this category have used the combination of a headset with a controller. As an outlier among the studies with interactive environments, there is an article where there was no headset and no controller. Yet, instead, the body's movement was used as a controller in a video game environment and a television screen to show the virtual environment.

3.3 Procedures, findings, and limitations

After looking at the similarity between the articles, it is now time to take a look at the procedure and the findings of each one to see how the technology impacted the learning process.

All the studies apart from but one that only had a post-test were conducted using a pre-test and a post-test, including a delayed-test. All the studies reported positive effects, even if mild, and increased results both in proficiency (language skills) and attitudes and easiness of the technologies used.

Nicolaidou et al. [9] showed that both the control group and the VR group were equally successful in this investigation. But it is essential to specify that the control group had the same application as the VR and the only difference was that they did not use a headset. In light of that information, one could speculate that a head-mounted VR headset per se does not offer any advantages, but the software does.

However, Repetto et al. [10] provide a study where the impact of immersive 360° on English vocabulary acquisition was analyzed. The findings highlighted that students who learned English by watching 360° films outperformed their peers who viewed non-immersive standard videos. The research also showed that the greater efficacy of 360° video instruction remained even after controlling for technology usage.

In the third article [11], the results show several advantages of using VR, mainly improved recall compared to the control group that used a flashcard system. There was also an increased advantage for AR in productive recall four days after the initial test. Finally, the qualitative survey and interview data indicated that participants believed that AR is effective and enjoyable for language learning.

As far as the duration of the experiment is concerned, Ma's study [12] lasted the whole semester, making it the most extended study in this review, and it is essential when compared to the other studies where there was little interaction with the technology. After a semester of instruction, the two groups were given a final English exam that included listening, speaking, grammar, reading comprehension, and writing. The data show that the experimental class's listening and speaking scores were much higher than those of the control class. Overall, the average writing score in the experimental class was slightly higher than that in the control class, which could be attributed to a solid conversational ability and excellent language sense. After the final exam, a survey on VR immersion teaching was conducted for the students in the experimental group. The findings reveal that the number of students who fully adapted to the VR immersion teaching approach was the same as the number of students who struggled to adapt, demonstrating that if it was to be used in education, adoption, and implementation of necessary knowledge should be increased. An interesting point about this study is that the experiment had two points of differentiation between the control group and the experimental one. First, constructivist pedagogy is combined with VR technology. Second, even though the results were quite positive, one cannot attribute them exclusively to technology.

Shadiev et al. [13] is a good research comparison to Reppeto's study since the main activity of this study is also 360° videos. In the beginning, questionnaires were administered to participants before the learning activity to assess proficiency in English. The

participants then engaged in learning exercises using a 360° video. After the experiment, the paired sample t-test showed significant differences between the pre and post-test after the VR lesson. The investigation yielded two crucial conclusions. The findings revealed that intercultural learning activities assisted by 360° video technology increased students' EFL ability. Finally, the students had a positive attitude toward the learning activities aided by 360° video technology, were happy with the technology, and planned to use it for learning in the future.

The case of the Kinect mediated experiment [14] is an interesting one because virtual reality is not defined by the use of a headset but by the use of a camera with body movement recognition. The Kinect allows the user to control the scene using body motions and vocal commands. Because of its advanced camera and speech recognition capabilities, Kinect lets users use their bodies as controllers in games and displays. A pre-test and post-test experimental designs were used in this investigation. The students who played a Kinect-based game learned more English military vocabulary than students who followed the usual curriculum and lesson design with a similar vocabulary corpus. The treatment (Kinect-based instruction vs. regular instruction) was the independent variable; the post-test performance was the dependent variable. In the qualitative portion of the study, students' perspectives were exposed to Kinect-based learning environments. The results revealed a substantial mean difference in student accomplishment in post-test scores in favor of the experimental group. Furthermore, the study investigated students' attitudes toward using gesture-based computer systems for instructional purposes. The findings revealed several critical factors to consider when using Kinect-based games for educational purposes, such as interaction and student motivation. Furthermore, it is vital to emphasize that language learning via a Kinect-based game made a kinesthetic contribution to students' accomplishments in terms of virtual embodiment learning. As long as students play the game, they can embody some of the gaming tools. Furthermore, the Kinect component of the study assisted students in gaining procedural expertise in language vocabulary learning.

In the last article, Wang et al.'s [15], participants were required to complete pre-tests on the topic of onomatopoeia. The pre-test results assessed the participants' cognitive knowledge of onomatopoeia before seeing the VR content. Each participant initially went through a calibration process for the VR eye tracker to record the correct viewing position in a separate study facility.

Participants entered a 3D VR Land theme park, which included five major amusement parks and a climate situation for mimicry and onomatopoeia themes: climate, speed, mood, rotation, animal sounds, and food temperature. Participants could freely see the 360-degree VR scenes, use the handheld controller to operate and pick facilities at their leisure, begin to engage with the facilities, and learn the onomatopoeia phrases associated with the facility. The VR session lasted roughly 15 minutes (3 minutes per amusement park). Following the VR trial, each participant immediately completed a post-test to assess the efficacy of the VR.

After seeing the VR simulation content of Japanese onomatopoeia, all participants' post-test scores were higher than those in the pre-test, according to their pre-test and post-test scores, whether they were in the high prior knowledge group or the low prior knowledge group. This difference is statistically significant. The findings indicate the

use of immersive content. It is important to note that the original article emphasized the eye-tracking possibilities of VR than studying the effect of VR on learning.

Table 2 below then provides a summary of these findings.

Table 2. Summary of the key findings

Technology used	Key Findings	Limitation
Application using head-mounted VR display (Oculus Rift)	Increase in vocabulary performance when learning a new language.	Convivence sampling, small sample size, short duration
360° videos wearing a card-board VR display	Superior efficacy of the 360° videos training having controlled for the technology usage	Small sample size, self-assessment of initial proficiency
Microsoft HoloLens, an augmented reality head-mounted display.	Increase in productive recall compared to control group.	Novelty effect of VR, small sample size
VR technology-based immersive virtual context teaching from the perspective of constructivism.	The experimental group scored higher in tests covering listening, speaking, grammar, reading, comprehension, and writing	Novelty effect of V
360-degree video technology used by the participants includes a camera to record and VR glasses	Significant improvements in post-test covering organizational communications, personal interactions, and practice.	Lack of control group
Microsoft Xbox Kinect. It enables the gamer to control the game through body movements and voice.	vocabulary achievement test scores, in favor of the experimental group, all the students expressed that Kinect-based gaming was better at improving their vocabulary than traditional materials	Novelty effect of VR
Through a VR eye tracker, participants' visual behaviors were tracked and recorded	All post-test scores were higher than those in the pre-test even those with low pre-test scores after they watched the VR simulation content of Japanese onomatopoeia	Small sample size

4 Discussion

As far as the research questions are concerned, i.e., which language skills benefited the most from the use of VR and/or AR technology and at what level of proficiency, the findings reveal that VR and AR can be used among students with zero previous knowledge up to an intermediate and advanced level of fluency. Therefore, virtual reality technology appears to be capable of improving the experience at all stages of the foreign language learning process. This is in line with other research studies on this topic, such as [16, 17]. However, teachers need more training in the use of VR so that they could use these VR and AR tools efficiently and thus make their students' learning process meaningful [16].

As far as the language skills are concerned, it is essential to mention that the VR experiments that were analyzed here were limited to the practice of input-related skills (listening and reading) since there is no way to speak or write in those experiments, and also that most tests (both pre and post-tests) involved only the input skills. Due to this

fact, it is reasonable to conclude that only listening and reading skills play a role in the acquisition of L2 using virtual reality. In light of the lack of VR experiences with options for interaction, it will be necessary for future research to test if virtual environments with chatbots might facilitate speaking and writing practice. In addition, in this process, the teacher's guidance will be necessary [18].

Regarding the research question dealing with the different definitions of VR and AR, the findings prove that there is no single definition. For some studies, VR meant just the use of the media through a head-mounted display. Yet, for other authors, it meant the use of virtual interactive environments not limited to the use of a head-mounted display, such as with the article that used the Kinect [6]. This lack of standardization makes comparisons challenging to establish since the only connection is the use of technology, making it possible to conclude the use of technology in the classroom.

Analyzing the recommendations given by the detected studies about suggestions for future research and their limitations, the following common points were identified:

- The size of the samples was small, and the duration of the data collection for the study was short.
- There was a limited interaction with the technology, usually limited to one session.
- The lack of delayed tests and the recency effect of the pre-test and post-test do not allow to claim definitively the effects of the technology in long-term learning gains on language learning.
- Moreover, most of the studies were solely based on quantitative data.
- The impact of the novelty effects on the studies given that VR/AR can be classified as new and exciting technology.

Even though not all articles had these problems, it is possible to see a trend in these studies in which future research will need to revisit the experimental methods and test the validity of the results with larger groups over a long period to prove what the result of virtual reality is.

For future experiments, it will be necessary to think about the scalability of this process and how to apply it to the classroom and to more realistic learning environments, where there is the interaction between students and teachers. Additionally, it is vital to question to what extent VR should be part of the language experience, should it be the primary vehicle for the acquisition, or should it be a complement or even just valuable for some context.

5 Conclusion

The main objective of this systematic review was to discuss the most significant aspects of current research, offering additional helpful insights into the use of VR/AR in L2 acquisition and indicating some key aspects that require the researcher's immediate attention as they become more widespread in the language learning space. After inspecting all the research on the topic, it seems that all virtual immersion technologies have a positive impact.

As it has been mentioned in the discussion, there is space for improvements in the research since most of the experiments did not have a large group or had post-test that were applied after a considerable length of time and therefore, even if the results all lead to a positive outcome, they should be treated with skepticism.

As demonstrated by the results of Nicolaidou et al. [9] that show virtually no difference between the control group with an app on a mobile device and the same app but on a headset display, one can conclude that generally there is a positive effect. Still, there seems to be a more substantial effect when an interactive experience accompanies the head-mounted display.

6 Acknowledgment

This research was supported by the SPEV project 2022, run at the Faculty of Informatics and Management, University of Hradec Kralove. The authors thank Klara Tlaskalova for her help with data processing.

7 References

[1] B. Klimova and M. Pikhart, "New Advances in Second Language Acquisition Methodology in Higher Education", *Education Sciences*, vol. 11, no. 3, p. 128, 2021. https://doi.org/10.3390/educsci11030128

[2] B. Klimova, M. Pikhart, A. Cierniak-Emerych and S. Dziuba, "A Qualitative Analysis of Students' Reflections on the Current Use of Digital Media in Foreign Language Classes", *Sustainability*, vol. 13, no. 16, p. 9082, 2021. https://doi.org/10.3390/su13169082

[3] B. Klimova, "Use of Virtual Reality in Non-Native Language Learning and Teaching", *Procedia Computer Science*, vol. 192, pp. 1385-1392, 2021. https://doi.org/10.1016/j.procs.2021.08.141

[4] D. Kamińska et al., "Virtual Reality and Its Applications in Education: Survey", *Information*, vol. 10, no. 10, p. 318, 2019. https://doi.org/10.3390/info10100318

[5] S. Oyelere, N. Bouali, R. Kaliisa, G. Obaido, A. Yunusa and E. Jimoh, "Exploring the trends of educational virtual reality games: a systematic review of empirical studies", *Smart Learning Environments*, vol. 7, no. 1, 2020. https://doi.org/10.1186/s40561-020-00142-7

[6] Y. Chen and C. Hsu, "Self-regulated mobile game-based English learning in a virtual reality environment", *Computers & Education*, vol. 154, p. 103910, 2020. https://doi.org/10.1016/j.compedu.2020.103910

[7] E. Childs et al., "An Overview of Enhancing Distance Learning Through Augmented and Virtual Reality Technologies", *arXiv.org*, 2022. [Online]. Available: https://arxiv.org/abs/2101.11000 [Accessed: 13- Apr- 2022].

[8] J. Davila Delgado, L. Oyedele, P. Demian and T. Beach, "A research agenda for augmented and virtual reality in architecture, engineering and construction", *Advanced Engineering Informatics*, vol. 45, p. 101122, 2020. https://doi.org/10.1016/j.aei.2020.101122

[9] I. Nicolaidou, P. Pissas and D. Boglou, "Comparing immersive Virtual Reality to mobile applications in foreign language learning in higher education: a quasi-experiment", *Interactive Learning Environments*, pp. 1-15, 2021. https://doi.org/10.1080/10494820.2020.1870504

[10] C. Repetto, A. Di Natale, D. Villani, S. Triberti, S. Germagnoli and G. Riva, "The use of immersive 360° videos for foreign language learning: a study on usage and efficacy among high-school students", *Interactive Learning Environments*, pp. 1-16, 2021. https://doi.org/10.1080/10494820.2020.1863234

[11] A. Ibrahim, B. Huynh, J. Downey, T. Hollerer, D. Chun and J. O'donovan, "ARbis Pictus: A Study of Vocabulary Learning with Augmented Reality", *IEEE Transactions on Visualization and Computer Graphics*, vol. 24, no. 11, pp. 2867-2874, 2018. https://doi.org/10.1109/TVCG.2018.2868568

[12] L. Ma, "An Immersive Context Teaching Method for College English Based on Artificial Intelligence and Machine Learning in Virtual Reality Technology", *Mobile Information Systems*, vol. 2021, pp. 1-7, 2021. https://doi.org/10.1155/2021/2637439

[13] R. Shadiev, J. Yu and W. Sintawati, "Exploring the Impact of Learning Activities Supported by 360-Degree Video Technology on Language Learning, Intercultural Communicative Competence Development, and Knowledge Sharing", *Frontiers in Psychology*, vol. 12, 2021. https://doi.org/10.3389/fpsyg.2021.766924

[14] M. Urun, H. Aksoy and R. Comez, "Supporting Foreign Language Vocabulary Learning Through Kinect-Based Gaming", *International Journal of Game-Based Learning*, vol. 7, no. 1, pp. 20-35, 2017. https://doi.org/10.4018/IJGBL.2017010102

[15] C. Wang, J. Hung and H. Chen, "How Prior Knowledge Affects Visual Attention of Japanese Mimicry and Onomatopoeia and Learning Outcomes: Evidence from Virtual Reality Eye Tracking", *Sustainability*, vol. 13, no. 19, p. 11058, 2021. https://doi.org/10.3390/su131911058

[16] K.-M. Chuah and M. K. Kabilan, "Teachers' Views on The Use of Chatbots to Support English Language Teaching in a Mobile Environment", *Int. J. Emerg. Technol. Learn.*, vol. 16, no. 20, pp. pp. 223–237, Oct. 2021. https://doi.org/10.3991/ijet.v16i20.24917

[17] E. Akman and R. Ã‡akır, "Pupils' Opinions on an Educational Virtual Reality Game in Terms of Flow Experience", *Int. J. Emerg. Technol. Learn.*, vol. 14, no. 15, pp. pp. 121–137, Aug. 2019. https://doi.org/10.3991/ijet.v14i15.10576

[18] K. Boettcher and A. Behr, "Using Virtual Reality for Teaching the Derivation of Conservation Laws in Fluid Mechanics", *Int. J. Eng. Ped.*, vol. 11, no. 4, pp. pp. 42–57, Jul. 2021. https://doi.org/10.3991/ijep.v11i4.20155

8 Authors

Francisco de Paula Forero Pataquiva is a Linguistics major at State University of Campinas.

Assoc. Prof. PhDr. Blanka Klímová, M.A., Ph.D. is head of the Department of Applied Linguistics at the Faculty of Informatics and Management of the University of Hradec Králové, Czech Republic, where she is mainly involved in teaching English for Specific Purposes and the use of ICT in language teaching.

Article submitted 2022-04-18. Resubmitted 2022-05-24. Final acceptance 2022-05-26. Final version published as submitted by the authors.