Enhancing Academic Speaking Skills: An Immersive Virtual World Approach

Elaine Hoter[1], Hannah Azulay[1], Manal Yazbak[2]

1 Hebrew University of Jerusalem
2 Sakhnin College

Abstract

In today's interconnected world, spoken English is a required skill on the academic level. Students are required to reach the B2 level of English according to the Common European Framework (CEFR) as a prerequisite for graduating and entrance to vocations. Students need to dedicate hours of practice in order to reach this level. To address this challenge, this study examines an attempt to teach advanced language skills using an innovative virtual world platform to create an immersive, authentic and engaging setting. The purpose of this pilot study is to find out whether immersive and experiential learning, along with teacher presence, can enhance students' confidence and motivation to engage in English and ultimately improve their level of aural and oral skills. This study employs a mixed-methods design based on an end-of-course questionnaire encompassing both open and closed questions (N=41), comprehensive course data for all participants, pre and post-oral and aural assessment scores from an external assessment company (N=28), participant observation throughout the course (N=120) and seven in-depth interviews. Participants reported that their speaking skills (fluency, pronunciation, vocabulary, and grammar), as well as their listening skills, had improved significantly, and this was corroborated by the quantitative data. The findings reveal that participants demonstrated heightened motivation and confidence to speak English through an avatar in a gamified environment. This study suggests that language immersion facilitated by teacher presence and virtual world technology has tremendous potential for students, in particular for those who lack confidence in speaking. This innovative approach can serve as a transformative method to enhance academic speaking skills, providing a more enjoyable and effective learning experience. Its implications for language learning are discussed.

Keywords: Language skills, Immersive learning, Teacher presence, Virtual world, Gamification, Motivation, Confidence.

Introduction

English is a global language (Melitz, 2016; Speak Now, 2023), and the increasing demand for fluency in English (Speak Now, 2023) has led to a wealth of research dedicated to finding more effective ways to teach English (Richards, 2008). In the past, academic English courses (EAP, English for Academic Purposes) were based primarily on reading skills. However, today, students in non-English speaking countries are required to achieve a high level in all four skills: listening, writing, reading and speaking. Many academic institutions, as well as employers in Europe, now require the B2 level in English (as stipulated by the European Union's Common European Framework of Reference for languages).

In response to this new requirement, academic institutions have had to rethink how they teach academic English in higher education. The speaking skill is particularly problematic because not only does it require adequate opportunities to speak in the language (Sun et al., 2017; Hakim & Rima, 2022), which a normal classroom situation cannot provide (Sun et al., 2017; Levak & Son, 2017), but also the speaking skill is generally accepted to contain four sub-skills: fluency, pronunciation, vocabulary and grammar (Astorga et al., 2011; Iwashita et al., 2008). In addition, there are personality considerations: extrovert/introvert as well as confidence and motivation issues (Liu, 2006).

This study looks at the potential for a virtual immersive environment as a solution for practising English to attain a high level of spoken English. The online immersive environment that is researched in this paper is called Englehaven, which was designed specifically for the purpose of B2 language instruction. It is considered a virtual immersive course due to its use of virtual world (VW) technology and avatars. 'Immersive', as used in this study, means when the participant is under the illusion that they are physically present in a different non-physical environment (Slater & Sanchez-Vives, 2014 and 2016; Freina & Ott, 2015). It is an instructor-led course where, each week, the material is released to the students. The course utilizes different methodologies incorporating technological elements, which help to maximize opportunities to speak in English. This study will mainly concentrate on the impact of a virtual environment on the speaking and listening skills of the participants.

Literature Review

Speaking and Listening

Different methods have been used throughout the years to analyze speech (Iwashita et al., 2008). The most prevalent method today divides speaking skills into four categories: fluency, pronunciation, vocabulary, and grammar (Saito et al., 2016).

Fluency

In linguistics, fluency is a term that refers to the flow of sounds, syllables, words, and phrases that are joined together when producing speech (Levelt, 1999). The most common method of assessing fluency focuses on three key aspects: speed of speech, breakdown fluency (how much the flow of speech is disrupted by repetitions or hesitations), and repair fluency (how much is corrected by the speaker) (Skehan, 2003). The present study aims at providing maximum practice in speaking with the goal of improving fluency in these three areas.

Pronunciation

Pronunciation can be defined as the phonological features of a language, including intonation, stress, and rhythm. A learner's ability to notice these phonological features reflects their internal processes and is crucial in EFL acquisition (Mackey et al., 2000). Prosodic features that also impact pronunciation, such as word stress, phrasal stress, pitch range, etc., can represent a challenge for language learners (Trofimovich & Baker, 2006; Kang, 2010). Assessing pronunciation along with other language areas allows a determination to be made of a speaker's intelligibility and to what extent their accent may or may not impede communication.

The opportunities presented in the present study for listening to new vocabulary and practising vocalizing them aim to aid in participants' pronunciation. Furthermore, learning prosody requires exposure to the language and practice, and in the present study, frequent repeated exposure to prosodic features aims at increasing speech comprehensibility.

Vocabulary

Different approaches to learning and teaching languages have developed over the years, each with a different perspective on teaching vocabulary (Schmitt & Schmitt. 2014). The vocabulary range to have academic and professional success in another language is established to be between 3000-5000 words (Nation, 1993; Chujo & Oghigian, 2009). This could be why recent EFL teaching practices are putting an emphasis on learning vocabulary (Laufer & Goldstein, 2004). Adequate vocabulary knowledge is important in EFL for comprehension and written and oral language (Schmitt, 2008). When vocabulary is learnt implicitly, it impacts a learner's breadth of vocabulary, how well each word is known and used, and correlates with reading comprehension (Stæhr, 2009).

A method for teaching spoken English is the teaching of language chunks, also known as formulaic sequences or collocations. Formulaic sequences are combinations of words with different meanings, used in spontaneous speech production in different contexts and are retrieved by the user as if they were a single word with a distinct meaning (Wood, 2006; Oberg, 2013). These chunks can include common expressions and proverbs, amongst others. The teaching

of formulaic sequences is one way of being able to improve oral proficiency in the EFL setting (Oberg, 2013). Studies have shown that when L2 learners use these formulaic sequences, they reduce the number of pauses in their speech (Oberg, 2013; Wood, 2006). Overall, previous research shows that formulaic sequences are not just chunks to be memorized; they allow the learner to produce automated expressions without the need to pause or use a string of unfamiliar words and, in addition, can improve confidence and reduce anxiety (Mohammadi & Enayati 2018). The present study focuses on repeated usage to target vocabulary in diverse contexts.

Grammar

Grammar is an important language skill for successful communicative competence and the ability to speak and use grammatical concepts effectively in real-life situations (Purpura, 2004; Long, 2016). As a result, its instruction has been the subject of many studies, where explicit and implicit instructional approaches have been debated (Murphy & Hastings, 2006). Implicit grammar teaching relies heavily on context (Ellis, Loewen & Erlam, 2006; Murphy & Hastings, 2006; Guci et al., 2021). Implicit instruction does not refer to the grammar rule directly; rather, students are exposed to the structure in context, such as requesting, apologizing and ordering. Only in a meaningful and understandable environment and scenes can the learners acquire English grammar and knowledge points naturally. The present study provides students with a wide range of scenarios and contexts with the aim of implicitly exposing the grammar concepts to students.

Listening Comprehension

Listening comprehension can be defined as an active process requiring a range of skills such as distinguishing between sounds, understanding words and grammar, interpreting intonation, and retaining information to be interpreted in context later (Astorga, 2011; Richards, 2008). The decoding process known as bottom-up processing includes several stages: firstly, the phonetic level, then the syntactic level, followed by a semantic level to decipher meaning, and finally, literal understanding (Astorga, 2011; Vandergrift, 2004). The top-down process, on the other hand, focuses on the listeners existing knowledge to guide their understanding of what they are listening to (Astorga, 2011; Yeldham, 2018). In other words, the listener may be able to understand what is being said on a topic based on previous knowledge despite not being familiar with all the words they hear. It is generally agreed that listeners adopt both processes and do not exclusively employ one of these approaches (Yeldham, 2018; Siegel & Siegel, 2015). The present study uses a combination of these strategies: the top-down approach in terms of tasks encouraging spontaneous speech on the VW and the bottom-up approach with the gamified tasks on the VW. In addition, the participants were exposed to different accents and dialects in authentic environments and situations.

The question arises whether other language areas can be improved if there is an improvement in listening comprehension. A study by Feyten (1991) concluded that there was, in fact, a connection between listening ability and other areas of foreign language acquisition. Astorga (2011) further concludes that improving listening skills has significant effects on various oral skills, such as vocabulary, pronunciation, and speech register.

The idea of the present study was to investigate an immersive environment where users learn collocations, vocabulary, and grammar in various topics in English. Students then practice what they have learnt in various modes: by themselves, using gamification incentives, pair work, group work, collaborative learning, and meeting with a teacher to get feedback on their speaking and listening skills. In addition, the present study scaffolds participants using formulaic sequences and lists of words as cues instead of free speech.

Motivation in Language Learning

The notable psycholinguist Zoltan Dorynei has characterized motivation as the reason why people decide to do something, how much effort they put into pursuing what they do and how long they are willing to carry it out (Dornyei, 2001). Dornyei classified L2 motivational constituents into seven broad dimensions visualized in Fig. 1. There is a correlation between intrinsically motivated language learners (i.e., learners who have an internal interest in the subject matter) and successful outcomes (Wu, 2011, Herlambang 2021). This is because performance can be maintained over longer periods of time since these learners are more willing to invest in the activity. This is in contrast with extrinsically motivated learners (i.e., those who learn the language for personal gain) who exhibit minimal effort (Wu, 2011). Higher levels of motivation relate to self-confidence, which appears to result in an increased willingness to speak in the L2 (Yashima et al., 2004). Since not all students can be intrinsically motivated to study a language, there has been a wealth of research strategies to assist with motivation in language learning (Dornyei, 2013; Yang et al., 2020; Chen et al., 2020; Yen et al., 2011; Chen, 2020; Chin et al., 2019; Ilter, 2009).
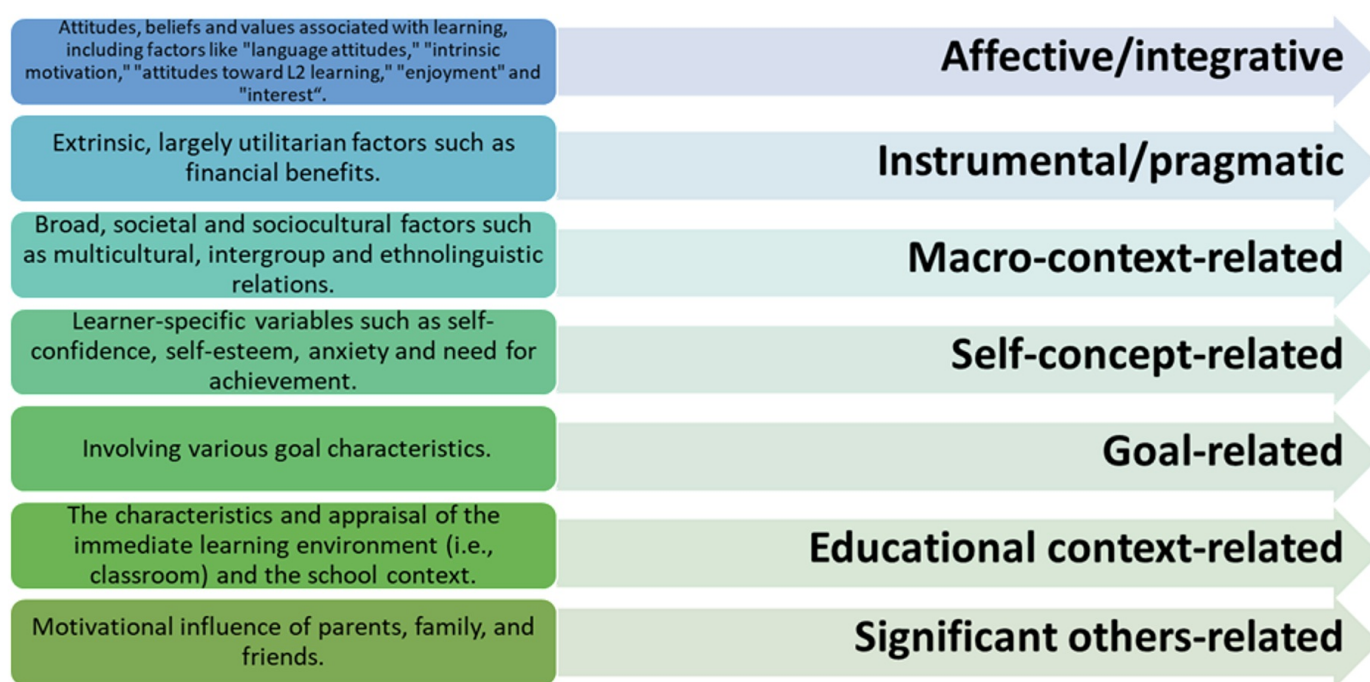
Figure 1. Dornyei's seven dimensions of L2 motivational constituents.

It is generally agreed that computer-assisted interventions facilitate students' positive learning attitudes toward learning a language (Chen, 2021). Therefore, the use of technology could empower students, giving them a sense of freedom and encouragement. Students feel invigorated by the visual and auditory elements, which trigger the senses and increase motivation. It can be concluded from previous studies that when technology is used in the learning process, students will be more interested in the lesson and its content, and their motivation will automatically increase (Chen et al., 2020; Elaish et al., 2019; Jiménez et al., 2020; Wu et al., 2011). The study will examine whether the use of an immersive VW has any effect on the motivation of students in their learning process.

Confidence in Language Learning

Students' confidence in learning a new language can be characterized by their willingness to communicate in the language (WTC) (Yashima et al., 2004; MacIntyre, 1994, Wang, 2010). WTC does seem to correlate with high levels of anxiety (Wang, 2010). Unlike motivation, which, under positive conditions, can rapidly increase, improvements in confidence can take time and practice (Wu, 2011). Foreign language anxiety (FLA) (MacIntyre, 1994) and foreign language classroom anxiety (FLCA) (Horwitz et al., 1986) are constructs that can affect language learning and language production and vary in degrees based upon higher and lower-level character traits (Dewaele, 2017).

It has been suggested that one way to assist language development and to achieve communicative confidence in the new language is for students to speak with native speakers (Yashima et al., 2004; Levak & Son, 2017), yet language learners who do not yet have the confidence to speak the language are less likely to initiate these interactions. New developments using AI give students the opportunities to enhance their language and communicative skills by providing AI chatbots as a way of practising and being exposed to the language (Hakim et al., 2022). This scenario provides a less intimidating way of practising with instant feedback, thus boosting users' confidence (Hakim et al., 2022).

Gamification

Gamification can loosely be defined as the incorporation of game-like activities into non-game contexts (Deterding et al., 2011). These activities have been noted to increase enjoyment levels and introduce a competitive factor (Jimenez-Sanchez & Gargallo, 2020). Yet even gamification elements that do not require external applications but rely on a gamified game environment, such as role-playing games, impact motivation and learning outcomes in second language learning (Piñero et al., 2022). Overall, a wealth of studies have examined the potential benefits of gamification and how it can impact EFL learning performance (Chen et al., 2019; Sanosi, 2018; Wu, 2018), the ways it assists with L2 motivation (Elaish et al., 2019; Wu, 2018) and reduce FLA in learners (Weissheimer et al., 2019; Zou et al., 2019). In the present study, computer technology has been used with the aim of allowing students to improve at their own pace and get instant feedback and with the goal of making an exciting, enjoyable learning experience.

Metaverse

Over the years, the metaverse has been referred to as artificial reality, virtual reality, and extended reality (Lee et al., 2021). However, it first came under discussion in the mid-1990s and referred to 3D virtual worlds where people interact with realistic virtual physical surroundings (Wigham et al., 2018; Lee et al., 2021; Mystakidis, 2022). In order to feel a part of this virtual environment, users are equipped with a digitally embodied representation of their physical-world self, which is called an avatar (Lee et al., 2021; Mystakidis, 2022; Hoter, 2017; Wigham et al., 2018). One feature of avatar usage in VR has been coined 'embodiment', or the feeling of ``being inside, having, and controlling a body" (Kilteni et al., 2012). Through both low and high-immersive technology, an embodiment can enable a person to feel that they are controlling a virtual body (Chen et al., 2021). Participants can also develop a deep connection to their avatar to the extent that they can feel pain in the real world when their avatar is inflicted with pain in the VW (Hoter & Nagar, 2021). Studies have shown that the use of an avatar different from oneself can contribute to behavioural changes, enabling people to identify with the characteristics of their avatar, thus increasing empathy and reducing negative stereotypes (Breves, 2020; Chen et al., 2021; Hoter & Shapira, 2022).

A wealth of research has attested to the use of VWs being more collaborative and engaging than traditional textbook usage (Yang et al., 2020; Hoter, 2017). In the area of English teaching, the use of virtual environments provides a promising solution for the lack of native-speaking environments in one's home country of instruction (Wang, 2017). According to Wang (2017), there is a correlation between learning outcomes and when learners feel a sense of immersion and presence in the VW. Presence is the extent to which users perceive the VW they are immersed in to be like their real world (Lee, 2004).

Research can attest to students feeling more present with the use of an avatar (Peterson, 2006) and, as a result, become less anxious speaking in English because interactions were not face-to-face (Wang et al., 2009; 2012) and their avatars' presence made them feel more confident to use the language (Yang et al., 2020). Students feel safe and relaxed, enabling them to be more willing to communicate (Lee & Hsieh, 2019). Second Life (SL) studies reveal better task execution and student engagement, encouraging more interactions, building confidence, and empowering the students via avatars (Chen & Kent, 2020; Levak & Son, 2017).

Several VR research studies have verified different challenges of using VR platforms. Certain experimental VR studies have attested to insignificant differences in learning outcomes when a VR platform was utilized (Hassenfeldt et al., 2020; Mackransky et al., 2019). Despite the motivating factor offered by immersive VR platforms, learning via this platform may overload and distract the learner, hence affecting opportunities for learning outcomes (Mackransky et al. 2019). The lack of access to technology, a low level of digital competence, as well as audio or visual disabilities, are also serious barriers that need addressing at pedagogical and institutional levels (UNESCO, 2022; Hoter & Nagar, 2023).

Virtual Reality Immersive Language Programs

Virtual reality (VR) in immersive learning has been described as a simulated environment where participants can interact with their surroundings and feel present in that environment (Yang et al., 2020; Lee & Wong, 2014; Makransky & Lilleholt, 2018). A number of studies have been carried out to show the effectiveness of an immersive VW on language skills. One such study is called Virtual Reality Life English (VRLE), which was developed specifically for the participants who were low-achieving high-school students to increase motivation in a disadvantaged English learning environment (Yang et al., 2020). To do this, they used various technological features such as high-quality graphics and sound and chatbots with native English dubbing. Their results showed that students were more engaged because they were less embarrassed about making mistakes, and because they had an all-round positive experience and felt their presence in the VW, they were more likely to participate vocally in English. Furthermore, when comparing their pre and post-oral test results, there were significant improvements. This shows that VR in language learning can facilitate oral language. Students felt immersed, and immersing themselves helped to improve their English.

There are additional gains when using VR in language programs. One VR language program found that participants enjoyed engaging with other avatars and would have wanted even more interaction with them, hence motivating participants to practice using the language (Yang, 2020). A VR language course with the aim of facilitating listening comprehension found positive outcomes because of the VW element (Levak & Son, 2017). This is because the course provided access to native speakers, cultural settings, exposure to online tools, instant messaging as one means of communication, and instant feedback. The purpose of this study is to find out what contribution the VW element of the course can have on speaking skills in the areas of fluency, vocabulary, grammar, pronunciation, and aural skills.

Teacher presence within the virtual world

Teacher presence within VWs can be defined as the ways in which the teacher interacts with the students in the VW and the extent to which they engage

them in the learning process (Ng, 2022). Ng (2022) purports the role of the teacher in VR to be more of a facilitator and motivator as opposed to being an instructor who merely transmits knowledge. Other studies also align with this theory and go on to further state that they even have the ability to reduce the digital divide and encourage students to fully express their learning (Garcia, 2010; Gandasegui, 2013). Some studies suggest that teacher presence within the world can contribute to the language improvement of the students.

In one study on the effects of SecondLife on speaking and listening skills, where an important feature of the way SL was used in this study was the presence of the teacher in scaffolding and giving directions and feedback to the students, notable improvements in speaking and listening skills were observed between pre and post-test for the t-test scores indicating that SL as an online VW has a significant effect on listening and speaking abilities (Shahri & Ashraf, 2016). In addition, students worked in pairs and in groups and had virtual face-to-face conversations. Another study showed that a teacher's presence in a VW helps students avoid a 'game-like mentality' when using SL for an EFL course and increases accountability (Wang et al., 2009). A more recent study conducted by Rojas and colleagues (2023) highlights the positive effects of teacher presence on VWs. Teachers on VWs can wield a profound influence on students when interacting with them on the VW through their dynamic interactions (Rojas et al., 2023).

The SEC Course

The Advanced Spoken English Language course (SEC), found on the EdX website (https://www.edx.org/), is a MOOC/VE which was designed with the intention of providing unlimited opportunities to speak in English using a wide variety of technological tools. The aim is that through these opportunities and completion of tasks using the different technologies, participants would improve their fluency, pronunciation, vocabulary, and grammar and, in turn, become more confident in communicating in English.

Englehaven Virtual Island

Students are required to spend time in the VW in a virtual town called 'Engelhaven', which was designed specifically to encourage speaking practice. Participants are assigned an avatar in the VW and are required to download and log in to the virtual island program (open-source SIM). This immersive environment contains weekly assignments that focus on various aspects of speech, such as commonly used proverbs and idioms. In addition, the island provides opportunities to learn spoken English in context. For instance, there is a unit based on real estate allowing participants to role-play being a persuasive realtor and a prospective home buyer, enhancing their language skills through experiential learning. In order to scaffold participants with their speaking skills, the VW leverages formulaic sequences as on-screen prompts for support.

Two evenings a week, an experienced English lecturer is present on the island to help students practice spoken English both individually and in groups and to do teacher tasks with the students, assess them and provide feedback to them when necessary. In addition to these interactions, weekly live technical support sessions and FAQ videos are available to address any technical queries or concerns students may have. There are many practice exercises that, when successfully completed, award you with 'Havens' – the currency in Englehaven. Each week, additional areas in the VW are made available with new materials and activities. Activities include opportunities to speak with NPCs (non-playing characters/robots), using cue cards (individual work), role plays and simulations (pair work), multiple choice of politeness responses in dialogue, word recognition, for example, near homophones, idioms, phrasal verbs, and gaming. Total practice time is up to the participant; it is estimated at 2 hours per week. All exercises completed by the participants are scored automatically and entered into a database that is accessible through each individual's profile.

Course content

The course consists of 10 units; as outlined here, each unit takes place in a different setting and covers different communicative functions. Both informal and formal uses of English are practised, and the units are pitched at progressively higher levels. Table 1 shows a breakdown of the different units in the course, showing which communicative function is being taught and practiced and in which context.

Table 1. Summary of the different units in the SEC course

Unit	Communicative purpose	Situation
1	Asking for information and being persuasive	Airport
2	Feedback and complaining British vs American English	Restaurant
3	Presenting oneself and introducing someone formally and informally	College campus
4	Appeasing and dealing with complaints	Municipality, town Hall
5	Describing situations & abstract concepts	Gallery, museum
6	Making excuses and promises	Recording studio
7	Persuading and asking for permission	Estate Agency, Bank, clinic
8	Agreeing and disagreeing: debating	University
9	Short speeches	Conference
10	Telling Jokes Improvizations	Comedy club

Figure 2. The Comedy Club (Unit 10)

Figure 3a. ATM Games
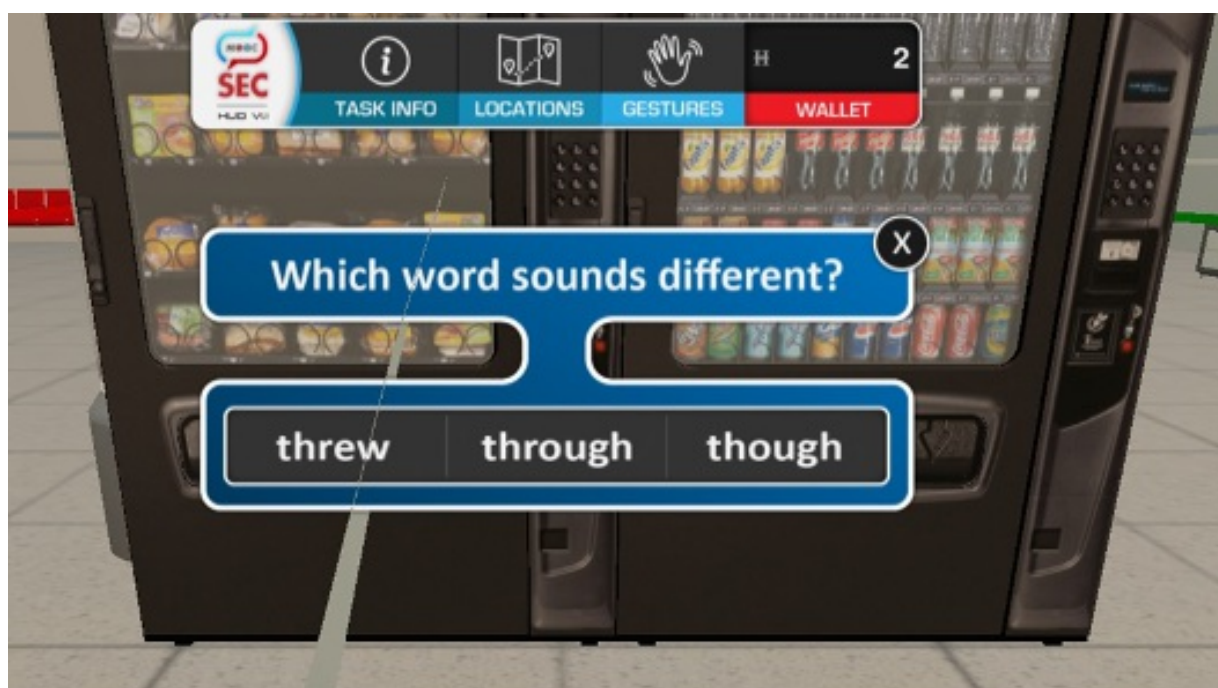
Figure 3b. ATM Games

Figure 4. Speaker's Corner (Unit 9) in Englehaven.

Methodology

Research Questions

This pilot study aimed to investigate the effects of an immersive VW on English as a foreign language (EFL) speaking and listening skills. Specifically, the study sought to examine whether the use of avatars and immersion in a VW would mitigate FLA and enhance participants' confidence. In addition, the study sought to explore the effectiveness of gamification elements incorporated into the course and identify the aspects of the course that contributed to greater improvement in speaking and listening skills. The Model for this virtual world included teacher presence. We wanted to see if this added to the course. The research questions guiding the study were:

1. To what extent did the participants' speaking and listening skills improve?

    1. Were there significant improvements from the pre and post-test?
    2. In which skills did participants improve?
    3. In which skills did the participants perceive improvement?

2. To what extent was there an improvement in participants' motivation and confidence to speak English?

    1. According to participants' perceptions?
    2. Which elements of the course contributed to participants' improvement in motivation and confidence?

3. What is the impact of teacher presence in the VW?

    1. on students' participation
    2. on students' perceived improvements

Hypotheses

1.

    1. There will be a significant improvement in the participants' speaking and listening skills from the pre-test to the post-test, as measured by a paired sample t-test.
    2. Participants will demonstrate improvement in specific speaking and listening skills, as observed and measured through both qualitative observations and a paired samples t-test.
    3. Participants will perceive an improvement in their speaking and listening skills, as evidenced by data obtained from interviews and surveys.

2.

    1. There will be a significant improvement in participants' motivation and confidence to speak English, according to their own perceptions, as gathered through interviews and surveys.
    2. The improvement in participants' motivation and confidence to speak English will be influenced by various course elements, including interviews, surveys, participant observation, course data (e.g., hours spent), and the points system.

3.

    1. The level of teacher presence in the VW will have a significant impact on students' participation in the course, as measured by course data and interviews.
    2. The extent of teacher presence in the VW will be positively associated with students' perceived improvements in their speaking and listening skills and accountability, as assessed through interviews and surveys.

Method

This pilot study employs a mixed-methods approach, encompassing both pre and post-tests as well as an extensive array of performance metrics to evaluate student progress. The quantitative aspect of the study entails a pre and post-external assessment test, which was subjected to analysis using AI algorithms, supplemented by a survey as outlined in this section. The qualitative elements involved data collected from open-ended questions and interviews. Forty-one participants answered a quasi-structured post-questionnaire, and there were seven in-depth interviews with participants after the completion of the course. During the reviewing and coding process of the interviews and open-ended questions in the questionnaire, themes were identified, and thematic categories were created to categorize the coded data. This also resulted in emerging sub-categories. After repeatedly reviewing the data organized under thematic categories, axial coding was applied to hone the themes and make connections between the themes and sub-categories, according to Corbin and Strauss (2014) and Robson (2016).

Participants

Two hundred students in higher education studying to be teachers from six culturally diverse educational colleges across Israel took part in the pilot advanced

speaking course. However, only 41 answered the post-course questionnaire, and 28 did the pre and post-test. Demographic information was collected using a questionnaire administered online, collecting data on age and gender. Participants were in the process of obtaining a B.Ed or M.Ed or were graduates taking the course for professional development. English is their second language (L2) or L3. A cross-section of participants was selected for the interviews in accordance with the following criteria: gender, age, cultural backgrounds, and undergraduate and graduate students.

Table 2. Demographic Characteristics of Survey and Interview Participants

		N	%
Gender	Female	32	78.04
	Male	9	21.95

Age	Mean	28.52
	(SD)	(9.84)

Materials

Materials include a background questionnaire, post-course survey, observations from the VW, feedback from teachers, data from the VW (time in the VW, grades for activities, completion rate, points for extra activities), pre-assessment and post-assessment of English (online through an outside company).

Reliability statistics

The data from the survey the students answered on completion of the course was analyzed for reliability. The overall internal consistency of this study's survey is high, with a Cronbach's Alpha coefficient of.928 based on 17 items. The corrected item-total correlations range from.219 to.885, which are all above the recommended minimum value of.20, indicating that all the items are measuring the same construct.

Individually, the items with the highest corrected item-total correlations (indicating the strongest association with the overall construct being measured) are "Fluency (Spoken level)" (.885), "Grammar" (.837), "Pronunciation" (.832), and "English expressions/idioms" (.824).

On the other hand, the items with the lowest corrected item-total correlations are "To what extent did the training unit and the videos prepare you for Englehaven?" (.313) and "How do you rate the technological support throughout the course?" (.329), indicating that these items may not be as closely related to the overall construct as the others.

The high Cronbach's Alpha coefficient and high corrected item-total correlations suggest that the survey is a reliable and valid measure of participants' perceptions of their English language improvement, as well as their experiences with the course and the teachers.

Table 3. Cronbach Alpha Results of Reliability from the Survey

Cronbach's Alpha	N of Items
.928	17

	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
Please rate how well you think your fluency has improved as a result of participating in this course.	.885	.917
Please rate how well you think your grammar has improved as a result of participating in this course.	.837	.918
Please rate how well you think your listening skills have improved as a result of participating in this course.	.761	.920
Please rate how well you think your confidence in speaking English has improved as a result of participating in this course.	.767	.920
Please rate how well you think your vocabulary has improved as a result of participating in this course.	.730	.921
Please rate how well you think your pronunciation has improved as a result of participating in this course.	.832	.918
Please rate how well you think your English expressions/idioms have improved as a result of participating in this course.	.824	.918
The teacher in the virtual world was friendly.	.462	.927
The teacher gave me useful feedback.	.624	.924
The teacher gave me sufficient time.	.624	.924
The teacher helped me gain confidence speaking	.664	.923
To what extent did the training unit and the videos prepare you for Englehaven?	.313	.930
To what extent do you feel that the course prepared you for the assessment at the end?	.665	.923
Would you take a similar online language course again in the future?	.219	.931
Would you recommend this course to others?	.620	.924
How do you rate the technological support throughout the course?	.329	.932
How useful was the MOOC Island site to track your progress?	.501	.927

Interviews

Since the make-up of the participant population in this study was diverse, individual interviews after the course were helpful in obtaining rich data about their perceptions and attitudes on the course. A cross-section of participants was selected for the interviews in accordance with the following criteria: gender, age, cultural backgrounds, and undergraduate and graduate students. The sessions were conducted in an informal atmosphere, as if students were having a casual conversation with a colleague. Since it is important to briefly describe the reasons for the study and the purpose of the interviews (Alsaawi, 2014), participants were briefed prior to the interview. The interviews were recorded and transcribed. Thematic analysis was chosen to interpret the interviews, and the interviews were indexed into categories (Corbin and Strauss; Robson, 2016). By using this method, emerging themes surfaced from the interviews, which aided in drawing valid conclusions. Below is information on each participant who was interviewed, along with their assigned anonymous alpha-numeric code, which will be used to refer to them in the results section:

- Participant 1 (P1): A 20-year-old female from the northern region of Israel, P1, is completing her undergraduate degree in education with a speciality in English.
- Participant 2 (P2): A 66-year-old female, P2, stands as a continuing student and lecturer at a teacher's college based in the central region. With her dual roles as both a learner and educator, P2 offered a unique perspective on her time in the VW.
- Participant 3 (P3): At 23 years old, P3, a female from the northern region finishing her undergraduate studies in education to be an English teacher, embodies the perspective of a younger generation in the study who is knowledgeable and familiar with technological tools. She connected and was enthusiastic about the innovative nature of the course. Her experiences shed light on how the VW resonates with modern learners and the difficulties she noticed her colleagues faced on the VW**.**
- Participant 4 (P4): A 19-year-old male from the southern region of Israel, P4 contributes a positive and enthusiastic voice to the study. He thrived on the personal attention received by the teacher and enjoyed the social interaction on the VW. He is also studying for his B.ed in Education with the intention of becoming an English teacher.
- Participant 5 (P5): A 34-year-old female residing in the southern region, P5 is finishing her B.ed. Her unique experiences on the VW shed light on how immersive learning strategies can cater to a diverse range of learner populations, including those of new mothers.
- Participant 6 (P6): P6, a 24-year-old female from the central region, who is completing her B.Ed. in education.
- Participant 7 (P7): A 40-year-old male based in the central region, P7's professional background in logistics gives a unique angle to the study since his motivation to do the course is for professional development.

Pre-test

Prior to the beginning of the course, students were given an opportunity to assess their oral fluency using the "speak now" software (speaknow.co.il) and scored in accordance with the CEFR levels of spoken English as mentioned above. This assessment includes security to prevent cheating and is filmed. Although the program assesses all 4 English skills (Speaking, Listening, Reading and Writing), we only assessed the students for Speaking and Listening. The SPEAK Assessment is a computer-based test of spoken English based on the principles of adaptive assessment, i.e. it begins at a general level, which determines the content and level of the items in the test itself. Spoken language is assessed for fluency, pronunciation, grammatical accuracy and vocabulary. The AI gives a grade from 0-6 for each of these areas, and there is an additional assessment for listening comprehension. The content of the test item relates to daily life, experiences, and expressing opinions (Speak Now, 2023).

Post-test

At the end of the 10-unit course, students are required to take the oral assessment via the SPEAK application. The test format is identical to the pre-test but with different items calibrated to make them equivalent to those on the pre-test.

Participant observation

Participant observation was utilized throughout the course, where data was gathered from a participant observer who was immersed in the VW. The observer spent time with the participants, carefully observing them while they performed tasks and inquired about their experiences in the VW. This approach enables the establishment of a strong connection with the participants, facilitating the observation of their authentic interactions during each session of the course. By observing and immersing oneself in the virtual community that will be created, an insider's perspective on the interactive and innovative elements of the VW was gained. This was an integral source of data, supporting both the qualitative and quantitative data and enhancing its reliability in the study.

Ethical Considerations

Upon enrollment in the SEC course, each student was informed of the purpose of the research, details about the tasks and activities (including audio and video recording of their performance), and assured that all work they submit will remain anonymous and confidential. The first question on the questionnaire requires the consent of the students to use the data from the form for research purposes.

All data was analyzed without reference to the student's identifying information. Every student received an alpha-numeric code to replace his/her personal data. Raw data (responses on tests, audio and video recordings, transcripts, etc.) was maintained on an external hard drive secured by a password. Data will not include personal identifying information from any participant. The research received ethical approval from an institute of higher education.

Results

The first question asks whether participants improved their speaking and listening skills and in what areas. In order to answer this question, we first carried out a Spearman's Rank Correlation Coefficient test. There was a high correlation between all the elements of the assessment, showing the reliability of the assessment. Table 4 presents the correlation coefficients and their significance levels between the variables of time spent in the VW (TimeVW), the number of tasks covered in the VW (CoverageVW), scores from the pre and post-SPEAK exam for Listening, Fluency, Pronunciation, Vocabulary and Grammar.

Table 4. Pre and post test correlations and significance levels in the different language areas.

		Correlations											
		Time VW	Coverage VW	Listening	Fluency	Pronunciation	Vocabulary	Grammar	Listening post	Fluency post	Pronunciation post	Vocabulary post	Grammar post
Time VW	R	1.000	.784**	-0.032	0.030	0.163	0.003	-0.044	-0.026	0.101	0.129	-0.116	0.000
	Sig. 2-tailed		0.000	0.819	0.828	0.239	0.981	0.751	0.898	0.625	0.529	0.572	0.999
	N	56	56	54	54	54	54	54	26	26	26	26	26
Coverage VW	R	.784**	1.000	0.053	-0.007	0.150	0.021	-0.034	0.109	0.112	0.197	-0.049	0.072
	Sig. 2-tailed	0.000		0.703	0.959	0.279	0.880	0.809	0.597	0.586	0.335	0.813	0.727
	N	56	56	54	54	54	54	54	26	26	26	26	26
Listening	R	-0.032	0.053	1.000	.613**	.579**	.635**	.555**	.463*	.523**	0.271	.389*	0.357
	Sig. 2-tailed	0.819	0.703		0.000	0.000	0.000	0.000	0.017	0.006	0.181	0.050	0.074
	N	54	54	54	54	54	54	54	26	26	26	26	26
Fluency	R	0.030	-0.007	.613**	1.000	.665**	.707**	.581**	.469*	.418*	.609**	.547**	.679**
	Sig. 2-tailed	0.828	0.959	0.000		0.000	0.000	0.000	0.016	0.033	0.001	0.004	0.000
	N	54	54	54	54	54	54	54	26	26	26	26	26
Pronunciation	R	0.163	0.150	.579**	.665**	1.000	.628**	.645**	.444*	.570**	.445*	.452*	.613**
	Sig. 2-tailed	0.239	0.279	0.000	0.000		0.000	0.000	0.023	0.002	0.023	0.021	0.001
	N	54	54	54	54	54	54	54	26	26	26	26	26
Vocabulary	R	0.003	0.021	.635**	.707**	.628**	1.000	.681**	.631**	.639**	.748**	.724**	.711**
	Sig. 2-tailed	0.981	0.880	0.000	0.000	0.000		0.000	0.001	0.000	0.000	0.000	0.000
	N	54	54	54	54	54	54	54	26	26	26	26	26
Grammar	R	-0.044	-0.034	.555**	.581**	.645**	.681**	1.000	0.385	0.232	0.322	.578**	.462*
	Sig. 2-tailed	0.751	0.809	0.000	0.000	0.000	0.000		0.052	0.254	0.109	0.002	0.017
	N	54	54	54	54	54	54	54	26	26	26	26	26
Listening post	R	-0.026	0.109	.463*	.469*	.444*	.631**	0.385	1.000	.696**	.490*	.793**	.587**
	Sig. 2-tailed	0.898	0.597	0.017	0.016	0.023	0.001	0.052		0.000	0.011	0.000	0.002
	N	26	26	26	26	26	26	26	26	26	26	26	26
Fluency post	R	0.101	0.112	.523**	.418*	.570**	.639**	0.232	.696**	1.000	.662**	.678**	.706**
	Sig. 2-tailed	0.625	0.586	0.006	0.033	0.002	0.000	0.254	0.000		0.000	0.000	0.000
	N	26	26	26	26	26	26	26	26	26	26	26	26
Pronunciation post	R	0.129	0.197	0.271	.609**	.445*	.748**	0.322	.490*	.662**	1.000	.594**	.746**
	Sig. 2-tailed	0.529	0.335	0.181	0.001	0.023	0.000	0.109	0.011	0.000		0.001	0.000
	N	26	26	26	26	26	26	26	26	26	26	26	26
Vocabulary post	R	-0.116	-0.049	.389*	.547**	.452*	.724**	.578**	.793**	.678**	.594**	1.000	.736**
	Sig. 2-tailed	0.572	0.813	0.050	0.004	0.021	0.000	0.002	0.000	0.000	0.001		0.000
	N	26	26	26	26	26	26	26	26	26	26	26	26
Grammar post	R	0.000	0.072	0.357	.679**	.613**	.711**	.462*	.587**	.706**	.746**	.736**	1.000
	Sig. 2-tailed	0.999	0.727	0.074	0.000	0.001	0.000	0.017	0.002	0.000	0.000	0.000	
	N	26	26	26	26	26	26	26	26	26	26	26	26

*R=Correlation Coefficient

There were also correlations involving the different language areas and components of the course. Listening proficiency was positively correlated with all the language areas, as well as with the tasks completed on the VW (CoverageVW). There was a strong positive correlation between listening skills and speaking fluency (r=0.613, p<0.001), which means that participants who had better listening comprehension tended to have better speaking fluency. There is also a strong positive correlation between speaking fluency (SpeakingFluency) and speaking pronunciation (r=0.665, p<0.001), which means that participants who had good fluency tended to have good pronunciation. Another strong positive correlation was between speaking pronunciation (SpeakingPronunciation) and speaking vocabulary (r=0.628, p<0.001), which means that students who had good pronunciation tended to have a good vocabulary.

Furthermore, there was a correlation in the pre-test between the components of the assessment: listening, fluency, vocabulary, pronunciation and grammar, meaning that these language areas are interconnected. T-tests were used to identify improvements between pre and post-assessments in the areas of fluency, pronunciation, vocabulary, grammar and listening, as shown in Table 5. For each pair, there are pre and post-test scores of an EFL skill area for the participants. For all pairs, the mean difference is negative, showing that participants scored higher on the post-test than on the pre-test. The t-statistics for fluency, pronunciation, and listening skills (pairs 1, 2, and 5) are negative and statistically significant at the 0.05 level, showing that the mean difference between pre and post-test scores is significant for these pairs. The t-statistics for vocabulary and grammar (pairs 3 and 4) are not statistically significant at the 0.05 level.

Table 5. Paired differences

		pre	post	difference	Lower	Upper	T	df	sig 2 tailed
Pair 1	Fluency - Fluency_post	-0.5962	0.8720	0.1710	-0.9484	-0.2439	-3.486	25	0.002
Pair 2	Pronunciation Pronunciation_p	-0.3846	0.9089	0.1783	-0.7517	-0.0175	-2.158	25	0.041
Pair 3	Vocabulary Vocabulary_p	-0.1538	0.6598	0.1294	-0.4204	0.1127	-1.189	25	0.246
Pair 4	Grammar - Grammar_p	-0.0962	0.8834	0.1732	-0.4530	0.2607	-0.555	25	0.584
Pair 5	Listening - Listening_post	-0.7692	1.2979	0.2545	-1.2935	-0.2450	-3.022	25	0.006

Table 5 presents the t-test results of the pre and post-tests for those students with over 37% coverage in the VW. The calculated t-statistic was less than the critical t-value for a two-tailed test, but it was greater than the critical t-value for a one-tailed test. Even though the values are weak to support a significant improvement in pre and post-test results for the two-tailed test, the one-tailed test at the 5% significance level can be used as evidence to support a significant improvement in post-test results for those participants who covered at least 37% of the VW elements. Thus attesting to the effectiveness of time spent in the VW and improvements in their language skills.

Table 6. T-Test: Two-Sample Assuming Equal Variances

	Variable 1	Variable 2
Mean	15.84210526	18.30555556
Variance	19.14035088	19.35702614
Observations	19	18
Pooled Variance	19.24559315	
Hypothesized Mean Difference	0	
df	35	
t Stat	-1.707223358	
P(T<=t) one-tail	0.04831929816	
t Critical one-tail	1.689572458	
P(T<=t) two-tail	0.09663859632	
t Critical two-tail	2.030107928	

just those with over 37% coverage

The first research question also asked in what areas the participants perceived improvement. In the post-course survey, the students were asked to rate their

English improvement in the following areas: fluency, grammar, listening, vocabulary, pronunciation and use of expressions and idioms. Table 7 shows the frequency of responses for each Likert scale response in the different language areas. Participants expressed relatively higher levels of agreement for the listening skill, with a notable number strongly agreeing that they made significant improvements in this area. Grammar and pronunciation, on the other hand, received mixed responses, suggesting diverse perceptions of proficiency. A relatively positive perception can be seen in the trend towards agreement for vocabulary and expressions/idioms.

Table 7. Likert scale responses on perceived improvements in the different language areas.

Skill	Strongly Disagree (1)	Disagree (2)	Neutral (3)	Agree (4)	Strongly Agree (5)
Fluency	7	5	15	11	4
Grammar	10	8	9	10	4
Listening	5	9	8	15	6
Vocabulary	4	7	12	15	5
Pronunciation	6	7	11	11	6
Expressions/Idioms	5	8	9	13	6

Table 8. Spearman's Correlation Coefficient Objective Measures

	Time VW	Listening	Fluency	Pronunciation	Vocabulary	Grammar	
Time VW	---	.08	.01	.02	.36*	.04	
Listening		---	.21	.32	.25	.37*	
Fluency			---	.11	.45*	.47*	
Pronunciation				---	.21	.38*	
Vocabulary					---	.13	
Grammar						---	

*p<.05, **p<.01.

From Table 8, which shows Spearman's correlation coefficient for objective measures, we can deduce that time spent in the VW was positively associated with better performance on campus and vocabulary improvement. Improvements in grammar were positively correlated with improvements in listening, fluency, and pronunciation, while pronunciation improvement was positively correlated with vocabulary improvement.

The Spearman's correlation table for subjective measures (see Table 11 below) shows there are significant positive correlations between student-perceived improvement, course evaluation, teacher evaluation, and course recommendation. Student perceived improvement is positively correlated with course evaluation (r=0.52, p<0.01), teacher evaluation (r=0.53, p<0.01), and course recommendation (r=0.56, p<0.001). Course recommendation is also positively correlated with teacher evaluation (r=0.32, p<0.05) and course evaluation (r=0.78, p<0.001). These findings suggest that students who reported higher levels of improvement in their English skills also tended to provide more positive evaluations of the course and teacher and would be more likely to recommend the course to others.

The second research question related to a perceived improvement in motivation and confidence and which elements of the course contributed to this. The students were asked to rate how their confidence to speak English had improved due to taking part in the VW, and the results are shown in Table 9; 66% felt their confidence had improved, moderately (29%) improved (37%) and substantially improved (12%) indicating they likely experienced a transformative change in their abilities.

Table 9. Likert scale responses for perceived confidence improvements.

	Please rate how well you think your confidence in using English in a variety of situations has improved as a result of participating in this course.	
	(1 is Not at all and 5 is Extremely well)	
Likert Scale Value	Number of responses	Percentage (%)
1	2	5
2	7	17
3	12	29
4	15	37
5	5	12

Open-ended survey questions and interview data were analyzed to find out whether students perceived improvements in their motivation and confidence. A major theme that emerged was an increase in motivation and confidence to speak and interact in English because of the VW experience. The sub-categories connected to this were inhibition, social aspect, success (leading to motivation and confidence), accessibility, immersion and presence (both teacher and student) and gamification. In contrast, motivation was decreased when there were technological issues.

Students identified with their avatar; furthermore, as one student said, 'having an avatar made me feel less inhibited' (P1). Many students attested to the fact that speaking through an avatar was less stressful for them as opposed to speaking in the real world, and they were able to overcome their anxiety about speaking in English. From the many comments about improved confidence and reduced anxiety, some excerpts taken from the interviews and questionnaire are: 'Because I had an avatar, I felt less nervous'(P6); 'I gained the confidence to make my voice heard' (S10); 'It will give me the confidence to talk and express myself'(S3); 'I gained the self-confidence I needed to actually talk in English' (P5); 'I knew how to speak very well, but I learned how to get more self-confidence when I am talking to someone in English' (P3). The atmosphere the VW created with the use of avatars enabled students to feel more confident in their abilities to speak in English and gave them the impetus to speak without any barriers.

There was also the social aspect of the course that also triggered motivation. Some participants worked with their friends from their home institution, which made the experience more enjoyable for them since they would organize to log in to the VW at the same time. Other participants made good friends during the VW from other institutions despite not keeping in contact after the course. Regarding this sub-category, some participants stated: 'Talking with other students and making friends helped with my speaking confidence' (S12). It was easier to make friends or find a partner in the VW: 'I enjoyed talking to other students I haven't met before and solving quizzes with them 'P3'; 'I enjoy the most in the VW because it is fun when my friends and I do the tasks together'(S6); 'The VW it was fun doing it with many students' (S13); 'I enjoy learning while playing and chatting with natives and other English speakers' (S17); 'I gained more experience talking with natives and other English speakers which was so fun' (P3).

Most significantly, the improvements that participants made in vocabulary and speaking contributed to a sense of success, which kept them motivated throughout the course. Some extracts present participants' feelings of success, 'I can now speak to others in English' (P3); 'It made me feel I could actually talk to people in English' (P5).

The sub-category of accessibility that emerged frequently in the qualitative data included the organization and user-friendly layout of each of the course elements, which made the course easy to navigate. A training session and a support system were put in place throughout the course that tried to assist students with technical issues. Videos were made to explain each unit and component of the course, which seemed to contribute to navigating around the course, and there was also a weekly Zoom technical support session where participants could get help. This increased the motivation of students to log in and complete the tasks. Another aspect of accessibility that students attested to was the flexible nature of the course; allowing them to enter the VW at their leisure was an important component that added to their motivation to study. Participants stated they would complete tasks whenever they had free time or would divide the work up for times when they would be free during the week. One participant said (P3), 'I would log in whenever I was bored'. One student said they wanted to log in because 'the assignments were fun and made me want to log in' (S5).

Another sub-category that emerged was embodiment, immersion and presence. Another reason that contributed to the confidence to speak in English was the feeling that participants were immersed in the VW, believing it was a real environment. Participants felt immersed to the extent they began to think in English since it felt natural for them to speak solely in English in the VW, 'I felt comfortable speaking in the VW' (S1). The VW gave participants endless opportunities to speak with other participants in group and pair work activities as well as tasks with the teachers, which students claimed facilitated confidence. The topic of immersion also surfaced within the qualitative data. The high-quality graphics that contributed to the ambience made the course fun and interesting for participants, 'The VW looked very modern and felt like a game' (P3). This led students to feel immersed in the VW, and this had a positive impact on their time

during the course. One participant also said, 'I felt immersed in the VW because I had many opportunities to speak and use the language in real-life scenarios' (P6).

Gamification was also a sub-category that emerged from the qualitative data, and when participants were asked about which elements increased their enjoyment, the gamified tasks were spoken of frequently. Students claimed the games were 'fun' and helped motivate them to invest more time to speak and practice in English. Two significant examples from the interviews that exemplify this are,

P4: 'The VW motivated me to do the tasks and speak to other participants. Also, the games in the VW, especially the one where we were at the restaurant. That was my favourite game because I learned how to order in a polite way. In the beginning, I was saying, 'I can get?', but the teacher told me, 'Say, 'I would like to order.'''

P3: 'The course was like a game, and the world is very modern, and I like these kinds of things. It reminds me of the game 'League of Legends' and games like that. It felt like I was playing them.'

Although participants commented on the value of the explanatory videos and extensive technical support offered by the course, the timing of the weekly technical support sessions wasn't convenient for everyone, 'There were weekly Zoom technical support sessions, but it was at a time when I was at college so I could not ask for help' (P5). There were technological issues which hampered participation, causing frustration and reduced motivation. This includes students' access to a computer with graphics and the infrastructure and good internet connection to fully participate in the course. For example, a computer that could not handle the software or a strong internet connection, 'my colleagues did not have the technology to participate like I could on the course' (P3). Another example of this and the resulting frustration is given by one participant (P5):

However, there were also some things I was not satisfied with. Tasks that should have taken 40 minutes, which would have been convenient as I have a small child at home, took me hours. In comparison to other students from my class, I spent the most time on the VW, and I did very well. But it took so much time because of all the bugs in the program, and it made me so frustrated. For example, I try to move my avatar, but it won't move. It was so frustrating because I had time, and I really wanted to work, but the bugs made it impossible. The technical issues took all the enjoyment away from the course.

This participant had to deal with a slow internet connection and an old computer at home, which would not allow her to function in the VW and could not function the avatar effectively. As a result, it took her longer to complete the assignments, which affected her levels of enjoyment of the course.

The final research question relates to the impact of teacher presence in the VW on student participation and perceived participant improvement in English. A close examination of the course data and the survey and interviews reveals the impact of teacher presence on student participation and perceived improvements. A special feature of the course was that in each unit, there was an assignment to do with a real teacher. A teacher was present in the VW twice a week to complete tasks and practice speaking with students. The teacher would then provide grades for some of the tasks and also verbal feedback to the students. The students could choose to log in to the VW during designated times when teachers would be present in the VW. The data shows that the highest levels of presence in the VW were during the teacher hours.

The students received assessments and feedback from the teachers for the teaching task. Students felt they benefited from the teacher modelling correct pronunciation and rich vocabulary in action, 'The teacher helped me to speak, and I learnt better pronunciation' (P7), and 'The tasks with the teachers were fun' (P2). In addition, many students stressed the importance of the teacher being patient and supportive in order to build a positive rapport with the participants, 'The teacher was supportive and always told us to take our time' (P4). Participants stated that they had a good relationship with the teachers on the VW, and some even said the teachers were 'the best I ever had' (P4). This also coincides with motivation and confidence since the good relationships built with the teachers increase students' motivation to log into the VW and confidence to speak in English.

The participants were asked four questions about the importance of teacher presence. Using a five-point Likert scale, they were asked to rate the teacher's role in the following domains: friendliness of the teacher, provision of useful feedback from the teacher, allocation of adequate time to the teacher, and assistance provided by the teacher in boosting confidence. For all questions, the majority of participants (47-59%) provided a rating of 4 (Agree) on the Likert scale, indicating that they generally agreed with the statements about the teacher's friendliness, feedback, sufficient time, and assistance with gaining confidence. A smaller percentage of participants (17-28%) chose a rating of 3 (Neutral) for each question. Responses indicating the lowest satisfaction (Likert value 1) are minimal across the questions, indicating that a small proportion of participants had negative experiences with the teacher on the VW. The percentage of participants who selected the highest rating (Likert value 5) varied between 6-22%, indicating that a smaller but still notable portion of participants had very positive perceptions of the teacher's attributes.

Table 10. Participant perceptions of the VW teacher.

Question:	The teacher was friendly		The teacher gave me useful feedback		The teacher gave me sufficient time		The teacher helped me gain confidence	
*Likert Value	N	%	N	%	N	%	N	%
1	0	0	1		0	0	1	3
2	2	6	3		3	8	6	17
3	4	11	6	17	6	17	6	17
4	7	19	8	22	10	28	9	25
5	23	64	18	50	17	47	14	39

Table 11 shows Spearman's correlation coefficient results between the subjective measures of students' sense of improvement, perception of the teacher, course evaluation and whether they would recommend the course. All the correlation values in the findings were statistically significant at the specified levels of significance (*$p < .05$, **$p < .01$, ***$p < .001$). Perceived improvement is moderately correlated with course recommendation ($r = .32^*$, $p < .05$) and teacher evaluation ($r = .78^{***}$, $p < .001$). Course recommendation is positively correlated with teacher evaluation ($r = .56^{***}$, $p < .001$). Course evaluation is moderately correlated with teacher evaluation ($r = .53^{**}$, $p < .01$).

Table 11. Spearman's Correlation Coefficient Subjective Measures

Perceived improvement	Perceived improvement	Teacher evaluation	Course evaluation	Recommendation
Teacher evaluation	---	.53**	.56***	.52**
Course evaluation		---	.12	.32*
Recommendation			---	.78***
				---

*$p<.05$, **$p<.01$, ***$p<.001$.

Discussion

The first research question dealt with the improvements in speaking and listening skills as perceived by the participants and reflected in the data. The results of this study show that students improved in all areas but most significantly in fluency, pronunciation and listening skills between the pre and post-test (see Table 4). By improving in fluency, the participants spoke faster with fewer hesitations and repetitions (Skehan, 2003). Pronunciation skills such as word stress, phrasal stress and pitch range improved (Trofimovich & Baker, 2006; Kang, 2010).

With regards to listening skills, the participants were more easily able to understand different accents and registers and improved their decoding skills (Richards, 2008; Siegel & Siegel, 2015; Vandergrift, 2004; Yeldham, 2018). The findings also suggest that improvements in listening skills can result in improvements in other language areas, corroborating the previous studies, which state that there is an interplay between advancements in all language areas (Feyten, 1991; Astorga, 2011). In addition, this study shows additional correlations; grammar improvements correlated with the other language areas, and there was a strong correlation between vocabulary acquisition and pronunciation, as shown in Table 4. Time on task significantly improved speaking and listening levels in all areas. Participants who covered at least 37% of the tasks in the VW (see Table 6) and practice made significant improvements in all language areas between the pre and post-test.

Regarding participants' perceptions of their improvement, a notable trend emerged. Participants felt they exhibited improvements in their listening skills, with a substantial number of participants expressing strong agreement about having achieved significant advancements within this area. In addition, participants perceived they had also improved in widening their vocabulary, improving their spoken ability and, as a result, felt they had become more fluent. Furthermore, as established by both the qualitative and the quantitative data, participants were satisfied with the style of learning (course evaluation) and connected with the VW technology, which reinforces the literature on positive learning outcomes and the use of VR technology (Yang et al., 2020; Chen et al., 2020; Yen et al., 2011; Chen, 2020; Chin et al., 2019; Ilter, 2009). The relationship between learning outcomes and perceived VW presence is also reinforced in this study, which echoes prior research (Yang et al., 2020; Lee & Wong, 2014; Makransky & Lilleholt, 2018). This suggests a close relationship between immersion in a

VW and enhanced speaking skills.

A number of interesting findings surfaced in the second research question relating to issues of motivation and confidence (Liu, 2006). Participants reported that their motivation improved due to their sense of feeling they were part of the VW. This sense of immersion confers with previous literature (Slater and Sanchez-Vives, 2014 and 2016; Freina and Ott, 2015). In addition, presence and embodiment felt while being on the VW can be attested to from their strong connection to their avatar and their feeling of presence in the VW conferring with the literature (Kilteni et al., 2012; Bailey & Bailenson, 2017; Nagar et al., 2021; Peterson, 2006). This sense of presence enabled them to feel part of a community since they had to interact with other participants through a variety of tasks. Their feeling of 'embodiment' enabled them to have enjoyable experiences on the VW, thereby increasing motivation, aligning with the findings from Chen and colleagues (2021). In addition, the auditory elements of the VW contributed to participants' heightened sense of presence, further bolstering motivation (Albahiri, 2020).

There was a profound impact from participants' sense of immersion, presence and embodiment in the VW on confidence. According to Dewaele (2017), producing language varies depending on individual personality, which includes their degree of confidence. Participants attributed this to the connection they had with their avatars, which is also in line with the literature on behavioural changes when avatars are used (Breves, 2020; Chen et al., 2021; Hoter & Shapira, 2022). The avatars created a 'buffer', reducing speaking-related anxiety in comparison to face-to-face interactions (FLCA) (Wang et al., 2009; 2012), effectively fostering a safer environment that boosted participation (Lee & Hsieh, 2019). This increase in confidence not only mirrors previous studies, notably Yang and colleagues (2020), but goes further to suggest that by boosting the student's confidence, all participants, and some for the first time, are able to practice and improve their spoken English without feeling intimidated.

Repeated simulations and role play in the VW were important in guiding students to use the taught expressions and idioms (formulaic sequences), thereby enriching their vocabulary and enhancing language proficiency (Wood, 2006; Oberg, 2013). There was a noticeable correlation between perceived vocabulary improvement and speaking success, which was accompanied by heightened motivation. This positive relationship, as in Mohammadi's research (2018), is attributed to participants being scaffolded by the incorporation of these predefined phrases, effectively mitigating their FLA.

Gamification elements within the VW acted as incentives, encouraging task completion and extrinsic motivation (Elaish et al., 2019; Wu, 2018). Participants reflected that these gamified aspects contributed to their levels of enjoyment, which encouraged them to keep returning to the VW to practice (Piñero et al., 2022), also reinforcing Chin and colleagues' assertion of VR's motivational impact.

Another salient pattern arising from participants' interviews and survey responses was their feelings towards the teacher's presence in the immersive VW and their perceived improvements in spoken ability. Not all VW's involve teacher presence. Participants who related positively to the presence of a teacher on the VW perceived an improvement in their spoken English, which has been attested to in both the quantitative and qualitative data. A few reasons for this seem to be the connection participants made with the teacher and the community built in the VW and the connection each made to their avatars (Kilteni et al., 2012). Since participants developed personal connections with their teacher, it can also be deduced that this aspect contributed to the participants' sense of accountability towards their learning and the course (Wang et al., 2009). This is catalyzed by the exchanges that took place between the participants and teacher, which reinforces the literature on the positive effects of the role and presence of the teacher in VWs (Ng, 2022; Shahri & Ashraf, 2016; Rojas et al., 2023). This evidence leads to the conclusion that the role of the teacher in a VW has the capacity to reduce the digital divide and encourage full language expression (Garcia, 2010; Gandasegui, 2013).

Conclusion

In conclusion, as well as reinforcing and producing theoretical insights into the use of VWs in education, the present study provides evidence specifically in support of improving all areas of EFL skills. It is also important to note how VW technology offers an alternative way of teaching online, combating the negative effects of two-dimensional online teaching systems (Mystakidis, 2022). The VW Englehaven world created combined both individual and collaborative learning through role plays and simulations as well as weekly assignments with teachers, thus providing support and authentic opportunities for speaking, which the traditional classroom lacks (Wang, 2017; Sun et al., 2017; Hakim and Rima, 2022; Levak & Son, 2017). The results of this study show the more students spent time in the VW connecting and speaking in English and completing the tasks, the better their results.

Implications have arisen from this study that go beyond validating previous theoretical insights and research in this area. It provides unique insights into how one academic spoken English course with an immersive VW feature could influence language skills and raise the level of participants in a short space of time. What can be established from the present study is the value of VWs in improving spoken English and listening skills and its potential to be an effective tool for language teachers or institutions seeking to improve students' language skills in a relatively short period of time.

Limitations

A few challenges were encountered in this study. Firstly, in terms of insufficient participation in both the pre and post-tests despite the large number of students who participated in the course. A larger corpus would be needed in future research to verify the results. The lack of participation could be attributed to initial technological difficulties. In addition, there was a steep learning curve for the initial setup of the VW, which resulted in some students facing difficulties in accessing and completing the tasks (Hoter & Nagar, 2023). Moreover, at this stage, the VW requires a computer as opposed to a phone or tablet, limiting participation primarily to those without access to this technology.

Recommendations

While the results of this study are promising, further research could determine the long-term effectiveness of VWs in language learning and explore other factors that contribute to their success on a larger scale. The corpus in this study was from teaching colleges across Israel. Further research could look at different populations to explore the varying impacts of virtual reality technology on language learning outcomes, taking into account factors such as age, language proficiency level, and cultural background. Despite existing technical support on the course, further research would provide extensive support to participants through the initial setup and the pre and post-testing process. On the basis of this research, we recommend Language Departments in Higher Education Institutions consider adding a VW component to their academic program.

References

- Alsaawi, A. (2014). A critical review of qualitative interviews. European Journal of Business and Social Sciences, 3(4).
- Astorga, E. D. (2011). The relationship between listening proficiency and speaking improvement in higher education: Considerations in assessing speaking and listening. Higher Learning Research Communications, 5(2).
- https://scholarworks.waldenu.edu/hlrc/vol5/iss2/5/
- Breves, P. (2020). Reducing outgroup bias through intergroup contact with non-playable video game characters. VR. PRESENCE: Virtual Augmented Reality 27, 257–273.
- Chen, J.C., & Kent, S. (2020). Task engagement, learner motivation and avatar identities of struggling English language learners in the 3D virtual world. System, 88, 102168.
- Chen, C. H., Hung, H. T., & Yeh, H. C. (2021). Virtual reality in problem-based learning contexts: Effects on the problem-solving performance, vocabulary acquisition and motivation of English language learners. Journal of Computer Assisted Learning, 37(3), 851-860.
- Chen, C. M., Liu, H., & Huang, H. B. (2019). Effects of a mobile game-based English vocabulary learning app on learners' perceptions and learning performance: A case study of Taiwanese EFL learners. ReCALL, 31(2), 170-188.
- Chujo, K., & Oghigian, K. (2009). How many words do you need to know to understand TOEIC, TOEFL & EIKEN? An examination of text coverage and high frequency vocabulary. Journal of Asia TEFL, 6(2).
- Dewaele, J. M., & Furnham, A. (2000). Personality and speech production: A pilot study of second language learners. Personality and Individual Differences, 28(2), 355-365.
- Deterding, S., Dixon, D., Khaled, R., & Nacke, L. (2011, September). From game design elements to gamefulness: defining" gamification". In Proceedings of the 15th international academic MindTrek conference: Envisioning future media environments (pp. 9-15).
- Dörnyei, Z. (2001). Motivational strategies in the language classroom (Vol. 10). Cambridge: Cambridge University Press.
- Elaish, M. M., Ghani, N. A., Shuib, L., & Al-Haiqi, A. (2019). Development of a mobile game application to boost students' motivation in learning English vocabulary. IEEE Access, 7, 13326-13337.
- Ellis, R., Loewen, S., & Erlam, R. (2006). Implicit and explicit corrective feedback and the acquisition of L2 grammar. Studies in Second Language Acquisition, 28(2), 339-368.
- Feyten, C. M. (1991). The power of listening ability: An overlooked dimension in language acquisition. The Modern Language Journal, 75(2), 173-180.
- Freina, L., & Ott, M. (2015, April). A literature review on immersive virtual reality in education: state of the art and perspectives. In The international scientific conference elearning and software for education (Vol. 1, No. 133, pp. 10-1007).
- Gandasegui, V. (2013). Entornos virtuales para el desarrollo de la educación inclusiva: Una mirada hacia el futuro desde el pasado de "Second Life" [Using

virtual worlds for the development of inclusive education: A glance at the future from the past of Second Life]. RELATEC. https://dehesa.unex.es/handle/10662/936

- Garcia, F. (2010). El uso de metaversos en el mundo educativo: Gestionando conocimiento en Second Life [Using metaverse in the educational world: Managing knowledge in Second Life]. Revista De Docencia Universitaria. 8(2), 147–159.

- Guci, R.I., Rochsantiningsih, D., & Sumardi, S. (2021). Focus on form and focus on forms in implicit grammar teaching strategy. Indonesian Journal of EFL and Linguistics, 6(1), 127-145.

- Hakim, R., & Rima, R. (2022). Chatting with AI Chatbots Applications to Improve English Communication Skill. Journal of English Language Studies, 7(1), 121-130.

- Hassenfeldt, C., Jacques, J., & Baggili, I. (2020). Exploring the Learning Efficacy of Digital Forensics Concepts and Bagging & Tagging of Digital Devices in Immersive Virtual Reality. Forensic Science International: Digital Investigation, 33(301011), 1-10.

- Herlambang, M. B., Cnossen, F., & Taatgen, N. A. (2021). The effects of intrinsic motivation on mental fatigue. PloS one, 16(1), e0243754.

- Horwitz, E. K., Horwitz, M. B. & Cope, J. (1986). Foreign language classroom anxiety. The Modern Language Journal, 70, 125-132.

- Hoter, E., & Shapira, N. (2022). Simulations in virtual worlds: improving intergroup relations and social proximity. Intercultural Education, 1-20.

- Hoter, E., & Nagar, I. (2023). The effects of a wheelchair simulation in a virtual world. Virtual Reality, 1-13.

- Hoter, E., & Nagar, I. (2023). Technical Support in Large Online Courses and MOOCs Using Virtual Worlds. The International Journal of Technologies in Learning, 30(2), 1.

- Hoter, E. (2017) A Model for Collaboration in Virtual Worlds Bringing Together Cultures in Conflict. In: Zaphiris P., Ioannou A. (eds) Learning and Collaboration Technologies. Technology in Education. LCT 2017. Lecture Notes in Computer Science, vol 10296. Springer, Cham.

- Iwashita, N., Brown, A., McNamara, T., & O'Hagan, S. (2008). Assessed levels of second language speaking proficiency: How distinct? Applied linguistics, 29(1), 24-49.

- Jiménez-Sánchez, M., & Gargallo-Camarillas, N. (2020). Gamification and Students' Motivation: Using Quizizz in the English as a Foreign Language (EFL) Classroom. Studia Universitatis Petru Maior. Philologia, (2), 143-157.

- Kang, O. (2010). Relative salience of suprasegmental features on judgments of L2 comprehensibility and accentedness. System, 38(2), 301-315.

- Kilteni, K., Groten, R., and Slater, M. (2012). The sense of embodiment in virtual reality. Presence Teleop. Virt. 21, 373–387.

- Laufer, B., & Goldstein, Z. (2004). Testing vocabulary knowledge: Size, strength, and computer adaptiveness. Language Learning, 54(3), 399–436.

- Lee, E. A. L., & Wong, K. W. (2014). Learning with desktop virtual reality: Low spatial ability learners are more positively affected. Computers & Education, 79, 49-58.

- Lee, J. S., & Hsieh, J. C. (2019). Affective variables and willingness to communicate of EFL learners in in-class, out-of-class, and digital contexts. System, 82, 63-73.

- Lee, L. H., Braud, T., Zhou, P., Wang, L., Xu, D., Lin, Z.,... & Hui, P. (2021). All one needs to know about metaverse: A complete survey on technological singularity, virtual ecosystem, and research agenda. arXiv preprint arXiv:2110.05352.

- Lee, K. M. (2004). Presence, explicated. Communication theory, 14(1), 27-50.

- Levak, N., & Son, J. B. (2017). Facilitating second language learners' listening comprehension with Second Life and Skype. ReCALL, 29(2), 200-218.

- Levelt, W. (1999). Producing spoken language. The neurocognition of language, 83-122.

- Liu, M. (2006). Anxiety in Chinese EFL students at different proficiency levels. System, 34(3), 301-316.

- Long, M.H. (2016). In defense of tasks and TBLT: Nonissues and real issues. Annual Review of Applied Linguistics, 36, 5–33.

- Makransky, G., Terkildsen, T. S., & Mayer, R. E. (2019). Adding immersive virtual reality to a science lab simulation causes more presence but less learning. Learning and instruction, 60, 225-236.

- Makransky, G., & Lilleholt, L. (2018). A structural equation modeling investigation of the emotional value of immersive virtual reality in education. Educational Technology Research and Development, 66(5), 1141-1164.

- Mackey, A., Gass, S., & McDonough, K. (2000). How do learners perceive interactional feedback? Studies in Second Language Acquisition, 22(4), 471-497.

- Melitz, J. (2016). English as a global language. In The Palgrave Handbook of Economics and Language (pp. 583-615). Palgrave Macmillan, London.

- Mohammadi, M., & Enayati, B. (2018). The Effects of Lexical Chunks Teaching on EFL Intermediate Learners' Speaking Fluency. International Journal of Instruction, 11(3), 179-192.

- Murphy, B., & Hastings, A. (2006). The utter hopelessness of explicit grammar teaching. The International Journal of Foreign Language Teaching, 2(2), 9-11.

- Mystakidis, S. (2022). Metaverse. Encyclopedia, 2(1), 486-497.

- Nagar, I., Hoter, E., & Hasler, B. S. (2021). Intergroup attitudes and interpersonal relationships in online contact between groups in conflict. Journal of Global Information Technology Management, 24(3), 208-223.
- Nation, I.S.P. (1993). Vocabulary size, growth, and use. The bilingual lexicon, 6, 115-134.
- Ng, D. T. K. (2022). What is the metaverse? Definitions, technologies and the community of inquiry. Australasian Journal of Educational Technology, 38(4), 190-205.
- Oberg, K. (2013). Formulaic sequences for improving oral fluency (Doctoral dissertation).
- Peterson, M. (2006). Learner interaction management in an avatar and chat-based virtual world. Computer Assisted Language Learning, 19(1), 79-103.
- Piñero Charlo, J. C., Belova, N., Quevedo Gutiérrez, E., Zapatera Llinares, A., Arboleya-García, E., Swacha, J.,... & Carmona-Medeiro, E. (2022). Preface for the Special Issue "Trends in Educational Gamification: Challenges and Learning Opportunities". Education Sciences, 12(3), 179.
- Purpura, J. (2004). Differing notions of 'grammar' for assessment. In Assessing Grammar (Cambridge Language Assessment, pp. 1-23). Cambridge: Cambridge University Press. doi:10.1017/CBO9780511733086.002
- Richards, J. C. (2008). Teaching listening and speaking (Vol. 35, No. 4). Cambridge: Cambridge University Press.
- Robson, C., & McCartan, K. (2016). Real-world research. Wiley Global Education.
- Rojas, E., Hülsmann, X., Estriegana, R., Rückert, F., & Garcia-Esteban, S. (2023). Students' Perception of Metaverses for Online Learning in Higher Education: Hype or Hope? Electronics, 12(8), 1867.
- Saito, K., Trofimovich, P., & Isaacs, T. (2016). Second language speech production: Investigating linguistic correlates of comprehensibility and accentedness for learners at different ability levels. Applied Psycholinguistics, 37(2), 217-240.
- Sanosi, A. B. (2018). The effect of Quizlet on vocabulary acquisition. Asian Journal of Education and e-learning, 6(4).
- Shahri, H. M., & Ashraf, H. (2016). On the Effect of Second Life (An Online Virtual World) on Pre-Intermediate Iranian EFL Learners Listening and Speaking Abilities. International Journal of English Language and Literature Studies, 5(1), 8-19.
- Schmitt, N. (2008). Instructed second language vocabulary learning. Language Teaching Research, 12(3), 329-363.
- Schmitt, N., & Schmitt, D. (2014). A reassessment of frequency and vocabulary size in L2 vocabulary teaching. Language Teaching, 47(4), pp?? (Note: The DOI was not provided.)
- Siegel, J., & Siegel, A. (2015). Getting to the bottom of L2 listening instruction: Making a case for bottom-up activities. Studies in Second Language Learning and Teaching, 5(4), 637-662.
- Skehan, P. (2003). Task-based instruction. Language Teaching, 36, 1–14.
- Slater, M., & Sanchez-Vives, M. V. (2016). Enhancing our lives with immersive virtual reality. Frontiers in Robotics and AI, 3, 74.
- Speak now (2023). The importance of testing candidates for English proficiency before hiring them for your company. Speak now
- https://speaknow.co/the-importance-of-testing-candidates-for-english-proficiency-before-hiring-them-for-your-company/
- Speak now (2023). English is a journey, not a destination: How Speaknow can help you make a start! Speak now
- https://speaknow.co/english-is-a-journey-not-a-destination-how-speaknow-can-help-you-make-a-start/
- Stæhr, L. S. (2009). Vocabulary knowledge and advanced listening comprehension in English as a foreign language. Studies in Second Language Acquisition, 31, 577-607.
- Sun, Z., Lin, C. H., You, J., Shen, H. J., Qi, S., & Luo, L. (2017). Improving the English-speaking skills of young learners through mobile social networking. Computer assisted language learning, 30(3-4), 304-324.
- Trofimovich, P., & Baker, W. (2006). Learning second language suprasegmentals: Effect of L2 experience on prosody and fluency characteristics of L2 speech. Studies in Second Language Acquisition, 28, 1-30.
- Vandergrift, L. (2004). Listening to learn or learning to listen? Annual Review of Applied Linguistics, 24, 3-25.
- Wang, C. X., Calandra, B., Hibbard, S. T., & McDowell Lefaiver, M. L. (2012). Learning effects of an experimental EFL program in Second Life. Educational Technology Research and Development, 60(5), 943-961.
- Wang, C. X., Song, H., Xia, F., & Yan, Q. (2009). Integrating Second Life into an EFL program: Students' perspectives. Journal of Educational Technology Development and Exchange (JETDE), 2(1), 1.
- Wang, Y.F., Petrina, S. and Feng, F. (2017), Virtual immersive language learning and gaming environment. Br J Educ Technol, 48: 431-450. https://doi.org/10.1111/bjet.12388
- Weissheimer, J., de Souza, J. G. M., Antunes, J. P. L., & de Souza Filho, N. S. (2019). Gamification and L2 vocabulary learning: The Vocabox experience in the languages without borders program. Revista Linguagem & Ensino, 22(4), 1136-1154.
- Wigham, C. R., Panichi, L., Nocchi, S., & Sadler, R. (2018). Interactions for language learning in and around virtual worlds. ReCALL, 30(2), 153-160.

- Wood, D. (2006). Uses and functions of formulaic sequences in second language speech: An exploration of the foundations of fluency. Canadian Modern Language Review, 63(1), 13-33.

- Wu, W.C.V., Yen, L.L., & Marek, M. (2011). Using online EFL interaction to increase confidence, motivation, and ability. Journal of Educational Technology & Society, 14(3), 118-129.

- Wu, T.T. (2018). Improving the effectiveness of English vocabulary review by integrating ARCS with mobile game-based learning. Journal of Computer Assisted Learning, 34(3), 315-323.

- Yang, F.-C. O., Lo, F.-Y. R., Hsieh, J. C., & Wu, W.-C. V. (2020). Facilitating Communicative Ability of EFL Learners via High-Immersion Virtual Reality. Journal of Educational Technology & Society, 23(1), 30–49.

- Yashima, T., Zenuk-Nishide, L., & Shimizu, K. (2004). The influence of attitudes and affect on willingness to communicate and second language communication. Language learning, 54(1), 119-152.

- Yeldham, M. (2018). L2 listening instruction: More bottom-up or more top-down? Journal of Asia TEFL, 15(3), 805.

- Zou, D., Huang, Y., & Xie, H. (2021). Digital game-based vocabulary learning: where are we and where are we going? Computer Assisted Language Learning, 34(5-6), 751-777.