The use of immersive 360° videos for foreign language learning: a study on usage and efficacy among high-school students

Claudia Repetto[a*], Anna Flavia Di Natale[b], Daniela Villani[a], Stefano Triberti[c,d], Serena Germagnoli[e], Giuseppe Riva[a,f]

[a] Dept. of Psychology, Università Cattolica del Sacro Cuore, Milan, Italy
[b] Dept. of Psychology, University of Milano-Bicocca, Milan, Italy
[c] Dept. of Oncology and Hemato-Oncology, University of Milan, Milan, Italy
[d] Applied Research Division for Cognitive and Psychological Science, European Institute of Oncology IRCCS, Milan, Italy
[e] SPAEE, Servizio di Psicologia dell'Apprendimento e dell'Educazione, Faculty of Psychology, Università Cattolica del Sacro Cuore, Milan, Italy
[f] Applied Technology for NeuroPsychology Laboratory, Istituto Auxologico Italiano, Milan, Italy
* corresponding author

Abstract

The large-scale diffusion of tablets and smartphones in the last decades offered new opportunities to broaden educational strategies. In particular, the Mobile-Assisted Language Learning (MALL) emerged as a promising approach leveraging the widespread availability of portable devices. In this study we investigated the feasibility and efficacy of a self-administered MALL homework training based on immersive 360° videos, compared to a training based on not-immersive standard videos showing the same visual content and enriched by an auditory description of the environment. The knowledge of target words was assessed before and after the training. In addition, students' attitudes toward the technology were assessed before the training. Results indicated that students with a more positive attitude toward technology watched the videos more than those with a less positive attitude. Furthermore, students who underwent the training with 360°videos learned more words than students belonging to the control group, even after controlling for the number of videos visualizations.

Keywords: immersive 360° videos; second language learning; virtual reality; embodied cognition, mobile-assisted language learning, acceptance

1. Introduction

The large-scale diffusion of personal computer, tablets and smartphones in the last decades has offered new opportunities to broaden educational strategies in formal and informal settings. The Computer-Assisted Language Learning (CALL) approach aims at delivering trainings for second language learning by means of computer-based programs such as online platforms (Usai, O Neil, & Newman, 2017), and videogames (Chen & Yang, 2013; Peterson, 2010). More recently, in the field of second language teaching another approach emerged, leveraging the widespread availability of portable devices, such as smartphones and media players: the Mobile-Assisted Language Learning (MALL). MALL has been distinguished from CALL in the use of personal portable devices (Kukulska-Hulme & Shield, 2008) emphasizing the concept of learning as an activity that one could perform “anytime and anyplace” (Geddes, 2004), thus extending definitely the learning environments beyond the boundaries of the classrooms. Importantly, within this approach students are in charge of their own learning process, and should feel more responsible of their achievements (Lai & Gu, 2011; Stockwell, 2008). The MALL approach is rapidly evolving, as technology implements new and more sophisticated features, improving graphics and interaction opportunities. Studies on MALL report a wide variety of learning protocols with different contents and aims (Duman, Orhon, & Gedik, 2014); despite the differences between studies, Viberg and Grönlund (2012), reviewing the literature on MALL from 2007 to 2012 found language learning benefits as well as positive attitudes towards the use of mobile technology for second language learning. Older studies focused on teaching vocabulary items delivered by means of short text messages (SMS) (Alemi, Sarab, & Lari, 2012; Lu, 2008; Thornton & Houser, 2005). Comparing the performance of students who learned the words presented by SMS (experimental group) with the performance of students who learned words on paper (control group), some of the studies reported that the experimental group outperformed the control group in the immediate recall (Lu, 2008; Thornton & Houser, 2005), whereas Alemi et al. (2012) described an advantage of learning by means of SMS only in the follow-up assessment. More recently, Liu and Chen (2015) used pictures to teach English as second language to Taiwan students. Specifically, students in the experimental group were presented with short phrases and were instructed to take pictures with their smartphones representing the phrases meaning in natural world settings. Authors reported better performances in the experimental group compared to the control group (involved in a phrase reading activity) only in the delayed test. Ducate and Lomicka (2013) used a completely different approach, not targeting any specific language ability, but rather examining the students' use of the different affordances offered by the mobile device, their utility in exposing to the target language and culture, and the students' attitudes toward the use of the mobile device as a learning tool. Their study was conducted along a semester, in which students were required to carry out a range of different activities in the target language with their mobile device: in-class tasks (i.e. searching information about cities, people and historical events, exploring newspapers headlines, watching YouTube videos); homework assignments (tweets on Twitter); out-of-class tasks (i.e. creating short videos, create photo collages, interviewing people). Authors analyzed both the products of the tasks and the surveys investigating the students' use of the device and opinions about it. They concluded that students took advantage of a variety of affordances (mainly social media and music for personal use; apps, dictionary and searching tools for academic purposes); in addition, students seemed to enjoy using the device as a learning tool, because it allowed a greater exposure to the target language and culture. It is worth noting that none of the above mentioned studies reported the effect sizes, thus hindering strong claims about training efficacy. Studies on MALL not only have analyzed students' usage patterns, but also students' perceptions (Hsu, 2013; Stockwell, 2008) as well as perceived usefulness and perceived compatibility between technology use and learning expectancies (Lai & Gu, 2011).

The most common theoretical model when investigating factors that predict the adoption of mobile learning is the Technology Acceptance Model (Davis, 1989) in its different versions (Huang, Huang, Huang, & Lin, 2012). This model recognizes two fundamental factors affecting people's attitudes toward technology and influencing at its turn the intention to use and the actual usage of the technology: perceived usefulness and perceived ease of use. Several studies investigated students' attitudes about MALL in different countries. Fujimoto (2012) surveyed Australian university students about the use of mobile phones in everyday life for non-educational, educational and language learning purposes; in addition learners' perception of the eventual use of mobile phones for language learning was investigated. Results indicated that most of the students had strong motivation to use their mobile devices for learning, especially if they had previous experiences with a didactic use of technology. Hsu (2013) made a cross-cultural analysis of university students' perception of MALL, including participants belonging to the Asian area. Findings showed that, even if students coming from different countries and with different cultural background had varying viewpoints about the possible affordances of MALL, they all agreed about the capabilities of MALL as constructivist tool. In addition, other studies investigating the Italian high school students' attitudes towards mobile technologies reported positive results (Cacciamani et al., 2018; Villani et al., 2018). However, it should be acknowledged that students are also aware of challenges and barriers associated to MALL (Dashtestani, 2016).

1.1. An embodied approach in MALL

Different methodological approaches are often reflected by different theoretical frameworks, even though not all the studies of MALL are clearly framed within a theoretical perspective. For instance, the research by Ducate and Lomicka (2013) was inspired by the Ecological Constructivism (Blyth, 2008): it focuses on different aspects of the learning process, including the student, the teacher, the environment, the technological devices, and the interactions between all these components. Sociocultural Theory of Mind (Vygotskiĭ & Cole, 1978) and Total Physical Response [TPR (Asher, 1988)] were the theoretical backgrounds in the research by Liu and Chen (2015). The first approach underlines the crucial contribution of society for the individual development, the second one is based on the coordination between language and action and proposes the physical activity as a mean to learn second language.

Another theoretical approach not yet applied in studies on MALL is the Embodied Cognition (EC) approach. According to EC, cognitive processes are grounded on sensory-motor experiences (Barsalou, 2008), therefore second language learning as well could be improved by trainings based on multimodal stimulations. Several studies demonstrated that adding a gesture to the word to be learned is beneficial for learning, both for concrete and abstract words (Macedonia & Klimesch, 2014; Mayer, Yildiz, Macedonia, & Von Kriegstein, 2015; Repetto, Pedroli, & Macedonia, 2017). Indeed, enriching words with motor information allows to create multimodal representations of the word meaning, and to retrieve the new word more accurately. However, enrichment can be accomplished not only by adding motor information, but also by means of multisensory stimulation: in this view, the word can be presented together with visual, auditory and motor inputs, creating a multisensory experience, deemed to be the optimal way our brain is tuned to work (Shams & Seitz, 2008).

Powerful tools to implement a multisensory, yet controllable, experience are Immersive Virtual Reality (IVR) systems. Fully immersive virtual environments are typically experienced by means of a Head-Mounted Display (HMD) or a cardboard (plastic or paper-made platform for a head mount), which allow users to exclude the physical world, being surrounded only by digital stimuli. Indeed, according to Slater and Sanchez-Vives (2016), immersion is a property entirely determined by technology and refers to the possibility to perceive the world through natural sensorimotor contingencies: for instance, wearing a HMD/cardboard I can use the movements of my head to explore the environment that surrounds me, exactly as I would do in real life situations. Together with the opportunity to explore the environment from a first person perspective, the visual perception of the digital content through natural movements affords a powerful "embodied experience" (Serino & Repetto, 2018). Indeed, recently Johnson-Glenberg and collaborators (2014) included immersion as one of the three components (being the other two the motoric engagement, i.e. the possibility to ambulate, and the gestural congruency mapping, i.e the coherence between the allowed gestures and the content to be learned) crucial for an educational technology to induce high degree of embodiment. Several studies in literature confirmed the impact of immersion on psychological dimensions and behavioral outcomes. Rupp and co-workers (Rupp et al., 2019), using 360° videos visualized with different devices, specifically investigated the effect of different degrees of immersion towards learning outcomes and subjective experience with university students. Authors found that more immersive devices induced greater positive affect and interest in learning the subject-matter and yielded better learning performances. Apart from educational contexts, immersive 360° videos demonstrated the capability to induce simple as long as complex emotions such as anger (Macedonio, Parsons, Digiuseppe, Weiderhold, & Rizzo, 2007), and awe (Chirico et al., 2017): this is very relevant also for learning, since together with perceptual features, also internal states evoked by the stimulus contribute to build some of the multiple representations the concept/word is stored with (Barsalou, 2008).

The potential use of IVR systems in education has attracted the interest of researchers and practitioners, and its efficacy for learning different school topics had been put under scrutiny, with mixed results (Bhattacharjee, Paul, Kim, & Karthigaikumar, 2018; Ekstrand et al., 2018; Parong & Mayer, 2018; Stepan et al., 2017); however, to the best of our knowledge, nobody has investigated yet the impact of IVR technology on second language learning within a MALL approach.

1.2. The present study

The main goal of the present research was to investigate the efficacy of a training for second language learning using a special kind of IVR, the immersive 360° videos, experienced by means of a smartphone combined with a cardboard. To pursue this goal, we prompted a group of Italian high school students to learn English words by means of a training based on immersive 360° videos and we compared their learning performance to that of another group of students who underwent a training based on not-immersive videos with fixed viewpoints. A quasi-experimental longitudinal study was designed to implement the trainings in ecological educational settings; more specifically, the trainings we proposed were not intended to replace the usual teaching method, but rather they were envisioned as a supportive tool for homework study.

Our research questions were the following:

1. does the group exposed to the immersive 360° videos training outperform the control group on English vocabulary learning?
2. does a more positive attitude towards technology positively affect its actual usage?

In addition, we aimed at measuring, in the group trained with 360° videos, the occurrence of cybersickness, a syndrome including symptoms such as headache, nausea, dizziness due to the conflicting information gathered from the visual, vestibular and proprioceptive sensory systems (Sharples, Cobb, Moody, & Wilson, 2008).

2. Method

2.1 Participants

The experiment was proposed to all the classes of eleventh-graders students attending the Technical Institute of Alessandria (3 classes, n= 70) and the Technical Institute of Milan (5 classes, n= 66), Italy. All the students, according to the teachers' assessment, belonged to the B1 level as defined by the Common European Framework of Reference for Languages (CEFRL).Thirty-two participants were excluded from analyses because either they presented intellectual disabilities, or learning disabilities or they were absent in one of the two in-class sessions. In the end, 104 participants were enrolled: sixty-three students from the Technical Institute of Alessandria were trained with immersive 360° videos (experimental group), while forty-one students from the Technical Institute in Milan were trained with not-immersive standard videos (control group). Since all of the participants were minors, parents signed an informed written consent for participating in the study. The study and the related consent form were approved by the University Ethic Committee.

2.2 Materials and contents

Twenty-one 360° videos were downloaded from the web. 360° videos are special types of movies shot with omnidirectional cameras that allow the collection of images from all around the space at the same time. The viewer can use different devices to watch these videos: in the desktop modality usually the image resembles a 3D sphere flattened onto a 2D rectangular frame, and the viewer can use the mouse or the keyboard arrows to control the view direction; in the immersive modality, the viewer must combine a cardboard with a smartphone (provided with a gyroscope), where the video is displayed. In this modality, the viewing experience is very close to reality: the user can control the viewing direction by means of the head movement in a realistic way (looking up one can see the sky/roof, looking down one can see the floor/ground, and if one wants to see what happens on the left/right, it is sufficient to turn the head accordingly, as in the real life situations). Of note, unlike other types of Virtual Reality, such as computer-generated environments, within 360° videos it is not possible to select the direction of the navigation, nor to interact with the elements of the environment.

All the videos were carefully examined by two researchers (CR and SG) with the purpose to exclude those that did not fit the experimental needs (e.g., videos with low quality of the images, content not suitable/appropriate, repetitive images, etc., were excluded). In the end, 10 videos were selected upon agreement, representing mainly natural landscapes (e.g. Africa, Venice, Japan, nature from the helicopter) and interior environments (e.g. house, hotel, car). Selected videos [already described by authors elsewhere (Repetto, Germagnoli, Triberti, & Riva, 2018)] were inspected scene by scene in order to identify the relevant elements included and the actions performed by the characters. Objects and verbs extracted were listed and became the target words to be learned in the second language (namely English, for Italian speakers). In agreement with the teachers involved in the study, words were included that belonged to the B1 and B2 levels of the CEFRL. On the whole, 148 items were used, being both high and low frequency words [Mean frequency = 168.87; SD = 316.48, according to COLFIS, Corpus e Lessico di Frequenza dell'Italiano Scritto (Bertinetto et al., 2005)]. High frequency words are supposed to be already known by the students, but still they were included to make the user feel confident with the task and to improve his/her self-efficacy. Furthermore, it was important that students could rely on some known words in order to understand the meaning of the new ones. In fact, for each video, a description had been created with the purpose to guide the students' attention and focus it towards the target elements named (e.g. "We're in Africa…We are nearby a wood. Look at the young man in a purple t-shirt…He's playing the guitar"). An English mother tongue woman audio recorded the descriptions, and the audio file had been added to the videos in the post-production phase, synchronizing the speech to the images displayed. Video duration ranged from 30 seconds to 1 minute and 30 seconds.

The same videos were edited to make them not-immersive standard videos, therefore traditional non-navigable videos in which the viewpoint is fixed. The editing procedure ensured that the images displayed in each moment were those named by the English voice. These control videos could be visualized on the smartphone, without any supplemental device.

2.3 Power analyses

In our design, participants were nested within condition and targets (in our case, word items) were crossed with condition. Each participant was in only one of the two conditions (360° videos vs. standard videos), but every target was judged under both conditions, albeit by different participants. For this reason, a priori power analyses was conducted using a specific tool to run power analyses with random targets and participants (Judd, Westfall & Kenny, 2016). The analyses indicated that a minimum sample of 75 participants was needed to detect a medium effect (d = .5) with a GLMM with 148 random targets.

2.4. Measures

To test the students' attitude towards the technology they were going to use, a brief ad hoc survey was administered, in which participants had to claim the degree of agreement with sentences presented, on a five-point Likert scale (ranging from completely disagree [1] to completely agree [5]). The survey was structured according to the Technology Acceptance Model (TAM) (Davis, Bagozzi, & Warshaw, 1989) and covering two main topics relevant for the present research: perceived ease of use and perceived utility of the technology (see Table 1 for the list of statements and reliability indexes). Both factors were scored calculating the mean items score.

To test the English vocabulary knowledge (T0) and the English vocabulary learning (T1), a paper and pencil test with the list of item (n=148) trained within the videos in the Italian version was proposed. The task was to translate the words in English. The bilingual translation tests (in T0 and T1) were scored assigning a value of 1 to the items correctly translated from L1 to L2, and 0 both to the items incorrectly translated and to the omissions.

2.5 Procedure

The training protocol had been designed in agreement with the teachers participating in the study, in order to fit the needs and constraints of the school. It consisted of four-weeks of self-training in which the students had to train themselves, individually and at home, with the immersive 360° videos or not-immersive standard videos provided (according to the experimental condition they belonged to). Before and after the self-training two sessions (T0 and T1) were held in the classrooms, during regular lessons, with one month of delay in between. An outline of the training protocol, including the measures collected, is represented in Figure 1.

Figure 1: Protocol of the study. The experimental group watched the 360° videos wearing the cardboard, while the control group watched at the not-immersive standard videos through the smartphone alone.

Figure 1: Protocol of the study. The experimental group watched the 360° videos wearing the cardboard, while the control group watched at the not-immersive standard videos through the smartphone alone.

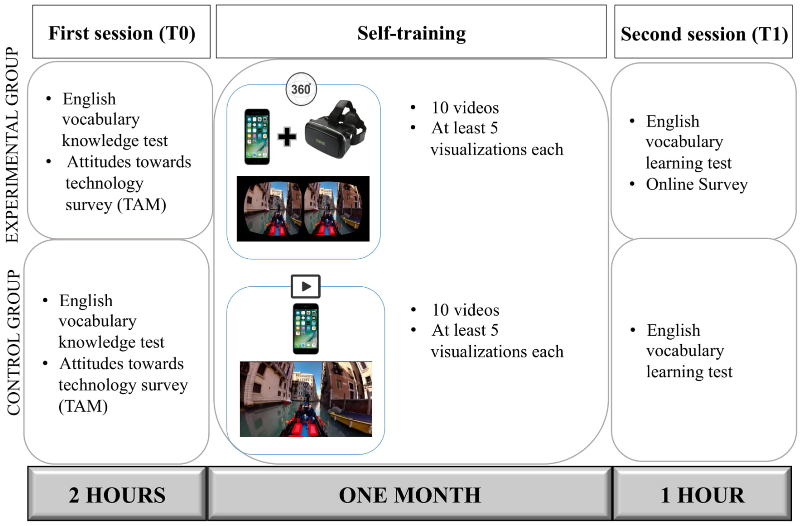

Figure 1: Protocol of the study. The experimental group watched the 360° videos wearing the cardboard, while the control group watched at the not-immersive standard videos through the smartphone alone.

2.5.1 First Session (T0)

The first session was dedicated to (a) the presentation of the project, (b) the collection of baseline measures related to English knowledge and students' acceptance about the use the two types of videos, and (c) the debriefing about the use of the technology. For the experimental group, each student had been provided with a personal Head-Mounted Display (iHarbort ® VR-G) to use for all the duration of the training. Researchers illustrated the use of the head-mounted display only in combination with students' personal smartphones and the videos' download and fruition by means of commercial freeware applications. When needed, a troubleshooting phase was arranged to fix issues arisen with the personal mobile devices. For the control group, students were shown how to download the videos and visualize them on the smartphone display.

At the end of the first session participants received a *vademecum* on the use of videos, aimed at maximizing the learning opportunities while reducing the possible risks connected with an unsuitable use of the technology. The document included the following recommendations: a. each video should be watched at least 5 times during the 4-weeks training; b. each visualization session should not exceed 5 minutes (no more than 3-4 videos for each session); c. at least two hours should pass between two consecutive sessions. Furthermore, participants were instructed to take note of the total number of visualizations.

2.5.2. Self-training

This phase comprised 4 weeks of training, in which students were supposed to follow the provided protocol. Researchers made themselves available to respond to any request of help or clarification by the students participating in the study.

2.5.3. Second session (T1)

During this session English learning had been assessed by the same bilingual translation test administered in the first session. Second, students were asked to report how many times they visualized each video during the self-training. Finally, only students of the experimental group compiled an online survey in which they saw the screenshots of the videos and had to rate the level of cyber-sickness experienced during each video's visualization, by moving a slide bar ranging from 0 to 100.

Table 1. The list of statements measuring students' attitudes for the experimental group; in brackets the alternative expression presented to the control group

TAM factors	Items	α
Perceived usefulness	I think that Virtual Reality (educational videos) will help me to improve my skills I think that Virtual Reality (educational videos) -based training will improve my English proficiency I think that Virtual Reality (educational videos) is useful	.67
Easiness of use	I think that learning to use Virtual Reality (educational videos) will be easy for me I think it will be easy for me becoming a Virtual Reality (educational videos) expert I think that Virtual Reality (educational videos) is easy to use	.83

2.6 Statistical analyses

To test the effectiveness of the immersive 360° videos training on English vocabulary learning it is recommended to consider that in language experiments there are two sources of random variability, the participants and the words (Clark, 1973; Coleman, 1964); therefore, we applied Linear Mixed Models to account for both random effects simultaneously (Janssen, 2012). Considering that our dependent variable (accuracy) was binomial (1= hit; 0= fail), the Generalized Linear Mixed Model (GLMM) approach has been selected. As predictors we considered one factor between subjects at two levels (Group: Experimental vs Control) one factor within subjects at two levels (Time: T0-pre and T1-post training), and one continuous variables being the number of visualizations (hereafter Visualizations). Following Barr et al's recommendation (Barr, Levy, Scheepers, & Tily, 2013), we first considered the maximal random model based on the design of the study. It included as fixed factors Time, Group and the two-way interaction Time x Group. Random intercepts were included for both participants (s) and words (w); in addition, the model included Visualizations as random factor across participants and words, and Time as random factor both across words and participants. The binary variables Group and Time were set so as the reference value against which the effect was calculated was 0 (which corresponds to the Control group and the pre-training time-point, respectively). The function representing the final model is the following (in brackets the random components):

$$y_{ws} = (u_{0s} + u_{w0} + u_{1s}*\text{Visualizations} + u_{2w}*\text{Visualizations} + u_{3w}*\text{Time} + u_{4s}*\text{Time}) + \gamma_{00} + b_1*\text{Group} + b_2*\text{Time} + b_3*\text{TimexGroup} + e_{ws}$$

To test whether the acceptance of technology affected its use, we built a multiple linear regression model with Visualizations as dependent variable and easiness of use and usefulness as predictors.

To understand, only at a descriptive level, to what extent the students belonging to the experimental group experienced cybersickness, we averaged by subject the scores attributed to the ten videos and then we computed descriptive statistics on that variable. Data were analyzed with SPSS software (Version 25, IBM Corp.)

3. Results

Table 2 summarizes descriptive statistics of the relevant variables.

Table 3 reports the parameters of the estimated GLMM. According to the model, accuracy varies random across words (u_{w0}), and participants (u_{0s}). Nevertheless, after controlling for the random factors, words were almost 2 times more likely to be correctly translated by students belonging to the experimental group than to the control group; similarly, words were almost 2 times more likely to be correctly translated in the post training assessment than in the first session assessment, indicating that, in general, students' knowledge of English words increased significantly after the training. The most interesting finding was the significant effect of the interaction Group X Time, which underlines that the probability to achieve a better performance after training is mediated by the group. More specifically, those who underwent the 360° videos training were one time and a half more accurate in the English proficiency test after training than those who underwent the not-immersive standard videos training (see Figure 2).

A multiple stepwise linear regression was calculated to predict videos visualization based on the usefulness and easiness students attributed to self-training. A significant regression equation was found [F(1,99)= 4.901, p<.05) with an R^2 of .047. In this model only the perceived easiness of use of the self-training predicted students' videos visualizations, while the perceived usefulness did not significantly impact on the number of video visualizations (see Table 4).

Looking at the descriptive data related to the cybersickness reported by participants of the experimental group, it is clear that few students answered this question (only 27 out of 63) but in general levels of cybersickness were very low (mean= 1.74).

Table 2. Descriptive statistics (N = number of valid cases; SE = standard error; Min. = minimum; Max. = maximum). English vocabulary scores are computed as the number of correct translated words by participants.

			N	Mean	SE	Min.	Max.
FIRST SESSION (T0)	**360° videos group**	**English vocabulary knowledge**	63	59.13	2.05	22	97
		Easiness of Use	62	3.56	0.12	1.33	5
		Perceived Usefulness	62	3.58	0.06	2	5
	2D- videos group	**English vocabulary knowledge**	41	47.34	2.13	27	82
		Easiness of Use	41	3.2	0.12	1.33	5
		Perceived Usefulness	41	3.16	0.1	2	5
SECOND SESSION (T1)	**360° videos group**	**English vocabulary learning**	63	77.7	2.86	33	122
		Visualizations	62	3.84	0.36	0	15
		Cybersickness	27	1.74	0.77	0	15
	2D- videos group	**English vocabulary learning**	41	56.22	2.96	24	110
		Visualizations	40	1.8	0.39	0	15

Table 3: Model summary of the GLMM. **= p<.005; ***= p< .001

	Fixed effects			*Random effects*	
				By subjects	By items
Parameters	**Exp (coefficient)**	**CI**	**t**	**σ²**	**σ²**
Intercept	0.24	0.14-0.4	-5.46***	0.16	1.83
Time (post training)	1.90	1.48-2.44	5.04***	0.21	0.15
Group (experimental)	1.97	1.29-3.00	3.13**	-	-
Time x Group	1.55	1.15-2.07	2.90**	-	-

Table 4. Stepwise multiple linear regressions between TAM factors and self-training visualizations (dependent variable)

TAM Factors	**Unstandardized Coefficients**		**Standardized Coefficients**		
	β	**Std. error**	**β**	**t**	**Sig.**
Easiness of use	.716	.323	.217	2.21	.029
Perceived Usefulness	.395	.474	.085	.833	.407

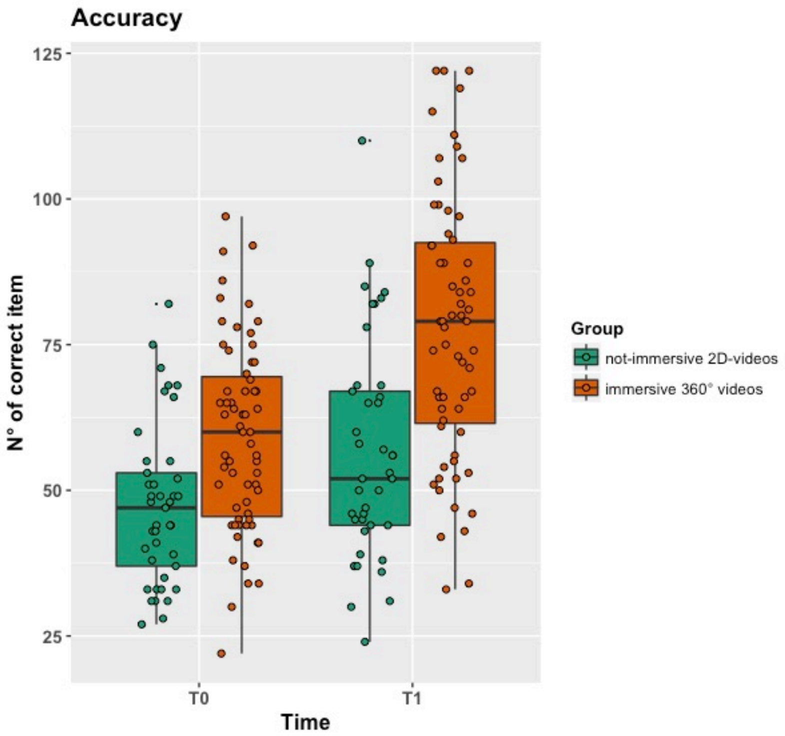

Figure 2: The Group by Time effect. Upper bars indicate the largest value within 1.5 times interquartile range (IQR) above 75th percentile; lower bars indicate the smallest value within 1.5 times interquartile range (IQR) below 25th percentile.

4. Discussion

In the present study, we compared the efficacy of a training based on immersive 360° videos with that of an alternative training based on not-immersive standard videos on second language learning. The trainings, proposed as homework assignments, were administered to eleventh-graders high school students of two Technical Institutes in Italy.

The first research question was related to the effectiveness of the immersive 360° self-training on English vocabulary learning. Findings underlined that students who learned English watching at the 360° videos outperformed their colleagues, who watched at not-immersive standard videos. The data indicated also that the superior efficacy of the 360° videos training remains even after having controlled for the technology usage.

Considering that the research protocol was designed to manipulate systematically the immersion (full immersion for 360° videos vs low immersion for standard videos), keeping constant the videos' content, we can maintain that the immersion was the key factor determining greater English learning outcomes. Our results are in line with previous studies that reported a beneficial impact of immersion on emotion induction (Chirico et al., 2017; Macedonio et al., 2007) and learning (Rupp et al., 2019). In our view, one possible determinant of this effect is the fact that immersive 360° videos offer an embodied experience. According to the taxonomy proposed by Johnson-Glenberg et al. (2014), the fourth, and highest, degree of embodiment can be achieved by educational technologies that allow users to walk, that provide the opportunity to use gestures linked to the learning content, and that are immersive, thus embedding the three core features of *motoric engagement*, *gestural congruency mapping* and *perception of immersion*, respectively. 360° videos cannot long to reach the fourth degree of embodiment, since they do not offer the opportunity to embed self-performed gestures inside the environment, but still can own, at least partially, two out of three of the components: they are fully immersive and, to some extent, they can provide the sense of locomotion as the first person perspective, the head-tracked changes of the point of view and the optic flow (Gibson, 1979) support the perception of self-motion.

Although this is the first study using 360° videos for second language learning, this technology has already been employed in educational settings for the study of different topics, typically STEM subjects. However, not all the learning protocols in which these type of videos had been used, reported positive findings in terms of learning outcomes. Ulrich and collaborators (Ulrich, Helms, Frandsen, & Rafn, 2019), for instance, administered a 360°video, a traditional video or a traditional lesson to physiotherapy students and stated that the spherical video was as effective as the standard one and of the traditional teaching in enhancing academic performance. More, traditional teaching turned out to be more appreciated than both types of video. Null results are described also by Chang et al. (2018), who included 360° videos in a platform to teach elementary students geomorphological knowledge. In this study, though, the comparison was made between two different learning approaches, both of which made use of the same 360° videos (i.e. an hands-on approach wherein students could build and edit the lesson content, versus a passive approach wherein students could only follow the lesson provided by means of the virtual reality-based platform). Thus, it is hard to compare our positive findings regarding the effectiveness of 360° videos as learning tools, with these negative or null findings: in the study by Ulrich (Ulrich et al., 2019), indeed, the content of the video was not created on purpose for the learning goal, but it was the mere video-recording, although spherical, of the traditional lesson; in the study by Chang et al. (2018), the object of the investigation was not the efficacy of the video per se but the comparison between two learning instruction strategies, therefore nothing can be said about the role of the video in this context. Only one study, to our knowledge, reported beneficial effects of this new technology on learning (Wu, Guo, Wang, Zeng, & Wu, 2019): in this experiment researchers proposed to a group of elementary students a learning activity based on the exploration of a 360° image illustrating a scientific concept; the control group, instead, viewed a video with common teaching explanations of the same scientific topic, combined with the exhibits of a science museum. Even though the 360 content was not conveyed immersively (the students saw the image on a tablet, using the translational and rotational movements of the device to explore the environment), those children belonging to the experimental group got better learning outcomes and greater problem solving abilities compared to their classmates belonging to the control group.

Concerning the role of students' attitudes on the number of videos visualizations, we found that perceived easiness of use but not perceived usefulness significantly contributed in predicting its actual usage during the self-training phase. Probably, the impact of the easiness of use on its actual usage could be explained in the light of the great confidence that adolescents, who are constantly surrounded by and immersed in new technologies, possess in using them (Palfrey & Gasser, 2011). Furthermore, the novelty effect of the self-training proposed could have enhanced the initial high-interest in performing the task, thus leading to higher visualizations (Stockwell & Hubbard, 2013). On the other hand, our result did not confirm previous studies about the influence of utility value on students' learning behaviors (Chiu & Wang, 2008) and the recognition of perceived usefulness as the most influential predictor of m-learning adoption (Liu, Li, & Carlsson, 2010). Probably, the fact that the training was proposed by the researchers and not by the teachers has penalized the perceived utility, thus not influencing the students' behaviors. The active engagement of the teachers could be a critical variable to sustain students' perceived usefulness towards this innovative self-training (Bozdoğan, 2015).

Finally, levels of cybersickness reported by students trained with 360° videos were very low, but the few answers collected did not allow considering this result a compelling evidence of the absence of side effects related to the use of technology.

5. Conclusions

The analysis of the present findings and of those reported in other studies using 360° video in educational settings, several recommendations can be drawn and some hints for future researches. In order to seek effectiveness, 360° videos should be created or edited as to fit the educational purpose they are supposed to achieve: in fact, the use of the tool as a mere way to provide distance learning (i.e. crossing the boundaries of space and time of a traditional lesson by allowing people to play the spherical video of the same lesson) does not yield any significant advantage over traditional teaching methods (Choi, Dailey-Hebert, & Estes, 2016). A second important aspect to consider it the students' acceptance of the technology: as evidenced by our data, the easiness of use predicted its subsequent actual usage but not the perceived usefulness. Based on this result, teachers interested in including new technologies in their learning activities should pay attention to the students' acceptance of that technology; they should also invest time and effort on one side in making the use of the devices and the related contents more friendly, on the other side in explaining the usefulness and the educational value of the technology-based learning strategy in order to support the continuance use of this technology (Chang, Yan, & Tseng, 2012; Mendoza, Carroll, & Stern, 2008).

As for the future perspectives, next steps could be to apply similar 360° videos trainings to in-school lessons, where more control can be carried out on the number of the videos administrations; in similar protocols the 360° videos training could be compared to other instructional activities, such as the traditional book-based language teaching.

Finally, considering the low expertise requested and the low budget needed, student could be trained to record themselves 360° videos to be used within learning approaches aimed to promote active students participation, such as the hands-on learning approach (Mullen et al., 2017), or the flipped classroom model (Roehl, Reddy, & Shannon, 2013). In conclusion, the present study demonstrated that 360° videos are effective technologies for second language learning and the 360° videos trainings are feasible in ecological learning settings, at least when they are proposed as supportive learning tools for homework study.

References

Alemi, M., Sarab, M. R. A., & Lari, Z. (2012). Successful learning of academic word list via MALL: Mobile assisted language learning. *International Education Studies*, *5*(6), 99–109. https://doi.org/10.5539/ies.v5n6p99

Asher, J. J. (1988). *Learning another language through actions: A teacher's guidebook* (3rd ed.). Los Gatos, CA: Sky Oaks.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278. https://doi.org/10.1016/j.jml.2012.11.001

Barsalou, L. W. (2008). Grounded cognition. *Annual Review of Psychology*, *59*, 617–645. https://doi.org/10.1146/annurev.psych.59.103006.093639

Bertinetto, P., Burani, C., Laudanna, A., Marconi, L., Ratti, C., Rolando, M., & Thornton, A. (2005). Corpus e Lessico di Frequenza dell'Italiano Scritto (CoLFIS). Retrieved from http://linguistica.sns.it/CoLFIS/Home.htm

Bhattacharjee, D., Paul, A., Kim, J. H., & Karthigaikumar, P. (2018). An immersive learning model using evolutionary learning. *Computers and Electrical Engineering*, *65*, 236–249. https://doi.org/10.1016/j.compeleceng.2017.08.023

Blyth, C. S. (2008). Research perspectives on online discourse and foreign language learning. In S. S. Magnan (Ed.), *Mediating Discourse Online* (pp. 47–72). Philadelphia: John Benjamins Pub.

Bozdoğan, D. (2015). MALL Revisited: Current Trends and Pedagogical Implications. *Procedia - Social and Behavioral Sciences*, *195*, 932–939. https://doi.org/10.1016/J.SBSPRO.2015.06.373

Cacciamani, S., Villani, D., Bonanomi, A., Carissoli, C., Olivari, M. G., Morganti, L., … Confalonieri, E. (2018). Factors Affecting Students' Acceptance of Tablet PCs: A Study in Italian High Schools. *Journal of Research on Technology in Education*, *50*(2), 120–133. https://doi.org/10.1080/15391523.2017.1409672

Chang, C., Yan, C., & Tseng, J. (2012). *Perceived convenience in an extended technology acceptance model: Mobile technology and English learning for college students*. *Australasian Journal of Educational Technology* (Vol. 28).

Chang, S., Hsu, T., Chen, Y., & Jong, M. S. Y. (2018). The effects of spherical video-based virtual reality implementation on students ' natural science learning effectiveness. *Interactive Learning Environments*, 1–15. https://doi.org/https://doi.org/10.1080/10494820.2018.1548490

Chen, H. J. H., & Yang, T. Y. C. (2013). The impact of adventure video games on foreign language learning and the perceptions of learners. *Interactive Learning Environments*, *21*(2), 129–141. https://doi.org/10.1080/10494820.2012.705851

Chirico, A., Cipresso, P., Yaden, D. B., Biassoni, F., Riva, G., & Gaggioli, A. (2017). Effectiveness of Immersive Videos in Inducing Awe: An Experimental Study. *Scientific Reports*, *7*(1), 1–11. https://doi.org/10.1038/s41598-017-01242-0

Chiu, C.-M., & Wang, E. T. G. (2008). Understanding Web-based learning continuance intention: The role of subjective task value. *Information & Management*, *45*(3), 194–201. https://doi.org/10.1016/J.IM.2008.02.003

Choi, D. H., Dailey-Hebert, A., & Estes, J. S. (Eds.). (2016). *Emerging tools and applications of virtual reality in education*. Hershey, PA: Information Science Reference.

Clark, H. H. (1973). The language-as-fixed-effect fallacy: A critique of language statistics in psychological research. *Journal of Verbal Learning and Verbal*

Behavior, *12*(4), 335–359. https://doi.org/10.1016/S0022-5371(73)80014-3
Coleman, E. B. (1964). Generalizing to a Language Population. *Psychological Reports*, *14*(1), 219–226. https://doi.org/10.2466/pr0.1964.14.1.219
Dashtestani, R. (2016). Moving bravely towards mobile learning: Iranian students0 use of mobile devices for learning English as a foreign language. *Computer Assisted Language Learning*, *29*(4), 815-832
Davis, F. D. (1989). Perceived Usefulness , Perceived Ease Of Use , And User Acceptance. *MIS Quarterly*, *13*(3), 319–339. https://doi.org/10.2307/249008
Davis, F. D., Bagozzi, R. P., & Warshaw, P. R. (1989). User acceptance of computer technology: a comparison of two theoretical models. *Management Science*, *35*(8), 982–1003. https://doi.org/10.1287/mnsc.35.8.982
Ducate, L., & Lomicka, L. (2013). Going mobile: Language learning with an ipod touch in intermediate french and german classes. *Foreign Language Annals*, *46*(3), 445–468. https://doi.org/10.1111/flan.12043
Duman, G., Orhon, G., & Gedik, N. (2014). Research trends in mobile assisted language learning from 2000 to 2012. *ReCALL*, *27*(2), 197–216. https://doi.org/10.1017/S0958344014000287
Ekstrand, C., Jamal, A., Nguyen, R., Kudryk, A., Mann, J., & Mendez, I. (2018). Immersive and interactive virtual reality to improve learning and retention of neuroanatomy in medical students: a randomized controlled study. *CMAJ Open*, *6*(1), E103–E109. https://doi.org/10.9778/cmajo.20170110
Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/BF03193146
Fujimoto, C. (2012). Perceptions of mobile language learning in Australia: How ready are learners to study on the move? *The Jalt Call Journal*, *8*(3), 165–195.
Geddes, S. J. (2004). Mobile Learning in the 21st century: benefit for learners. *Knowledge Tree E-Journal*, *30*(3), 214–228.
Gibson, J. J. (1979). *The ecological approach to visual perception*. Boston: Houghton Mifflin.
Hsu, L. (2013). English as a foreign language learners' perception of mobile assisted language learning: a cross-national study. *Computer Assisted Language Learning*, *26*(3), 197–213. https://doi.org/10.1080/09588221.2011.649485
Huang, Y.-M., Huang, Y.-M., Huang, S.-H., & Lin, Y.-T. (2012). A ubiquitous English vocabulary learning system: Evidence of active/passive attitudes vs. usefulness/ease-of-use. *Computers & Education*, *58*(1), 273–282. https://doi.org/10.1016/J.COMPEDU.2011.08.008
Janssen, D. P. (2012). Twice random, once mixed: Applying mixed models to simultaneously analyze random effects of language and participants. *Behavior Research Methods*, *44*(1), 232–247. https://doi.org/10.3758/s13428-011-0145-1
Johnson-Glenberg, M. C., Birchfield, D. A., Tolentino, L., & Koziupa, T. (2014). Collaborative embodied learning in mixed reality motion-capture environments: Two science studies. *Journal of Educational Psychology*, *106*(1), 86–104. https://doi.org/10.1037/a0034008
Judd, C. M., Westfall, J., & Kenny, D. A. (2017). Experiments with more than one random factor: Designs, analytic models, and statistical power. *Annual review of psychology*, *68*, 601-625.
Kukulska-Hulme, A., & Shield, L. (2008). An overview of mobile assisted language learning: From content delivery to supported collaboration and interaction.

ReCALL, *20*(03), 271–289. https://doi.org/10.1017/S0958344008000335
Lai, C., & Gu, M. (2011). Self-regulated out-of-class language learning with technology. *Computer Assisted Language Learning*, *24*(4), 317–335. https://doi.org/10.1080/09588221.2011.568417
Liu, P., & Chen, C. (2015). Learning English through actions: a study of mobile-assisted language learning. *Interactive Learning Environments*, *23*(2), 158–171. https://doi.org/10.1080/10494820.2014.959976
Liu, Y., Li, H., & Carlsson, C. (2010). Factors driving the adoption of m-learning: An empirical study. *Computers & Education*, *55*(3), 1211–1219. https://doi.org/10.1016/J.COMPEDU.2010.05.018
Lu, M. (2008). Effectiveness of vocabulary learning via mobile phone. *Journal of Computer Assisted Learning*, *24*(6), 515–525. https://doi.org/10.1111/j.1365-2729.2008.00289.x
Macedonia, M., & Klimesch, W. (2014). Long-Term Effects of Gestures on Memory for Foreign Language Words Trained in the Classroom. *Mind, Brain, and Education*, *8*(2), 74–88. https://doi.org/10.1111/mbe.12047
Macedonio, M. F., Parsons, T. D., Digiuseppe, R. A., Weiderhold, B. A., & Rizzo, A. A. (2007). Immersiveness and Physiological Arousal within Panoramic Video-Based Virtual Reality. *CyberPsychology & Behavior*, *10*(4), 508–515. https://doi.org/10.1089/cpb.2007.9997
Mayer, K. M., Yildiz, I. B., Macedonia, M., & Von Kriegstein, K. (2015). Visual and motor cortices differentially support the translation of foreign language words. *Current Biology*, *25*(4), 530–535. https://doi.org/10.1016/j.cub.2014.11.068
Mendoza, A., Carroll, J., & Stern, L. (2008). Influences on Continued Use of an Information System: A Longitudinal Study. In *Proceedings of 16th European Conference on Information Systems* (pp. 985–996). Galway, Ireland.
Mullen, J., Byun, C., Gadepally, V., Samsi, S., Reuther, A., & Kepner, J. (2017). Learning by doing, High Performance Computing education in the MOOC era. *Journal of Parallel and Distributed Computing*, *105*, 105–115. https://doi.org/10.1016/J.JPDC.2017.01.015
Palfrey, J. G., & Gasser, U. (2011). *Born digital : understanding the first generation of digital natives*. New York: Readhowyouwant.com.
Parong, J., & Mayer, R. (2018). Learning Science in Immersive Virtual Reality Article in Journal of Educational Psychology. *Journal of Educational Psychology*, *110*(6), 785–797. https://doi.org/10.1037/edu0000241
Peterson, M. (2010). Computerized Games and Simulations in Computer-Assisted Language Learning: A Meta-Analysis of Research. *Simulation & Gaming*, *41*(1), 72–93. https://doi.org/10.1177/1046878109355684
Repetto, C., Germagnoli, S., Triberti, S., & Riva, G. (2018). Learning into the Wild: A Protocol for the Use of 360° Video for Foreign Language Learning. *Lecture Notes of the Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering, LNICST*, *253*, 56–63. https://doi.org/10.1007/978-3-030-01093-5_8
Repetto, C., Pedroli, E., & Macedonia, M. (2017). Enrichment Effects of Gestures and Pictures on Abstract Words in a Second Language. *Frontiers in Psychology*, *8*, 2136. https://doi.org/10.3389/fpsyg.2017.02136
Roehl, A., Reddy, S. L., & Shannon, G. J. (2013). The flipped classroom: An opportunity to engage millennial students through active learning strategies. *Journal of Family & Consumer Sciences*, *105*(2), 44–49.
Rupp, M. A., Odette, K. L., Kozachuk, J., Michaelis, J. R., Smither, J. A., & McConnell, D. S. (2019). Investigating learning outcomes and subjective

experiences in 360-degree videos. *Computers and Education*, *128*, 256–268. https://doi.org/10.1016/j.compedu.2018.09.015
Serino, S., & Repetto, C. (2018). New Trends in Episodic Memory Assessment: Immersive 360° Ecological Videos. *Frontiers in Psychology*, *9*, 1878. https://doi.org/10.3389/fpsyg.2018.01878
Shams, L., & Seitz, A. R. (2008). Benefits of multisensory learning. *Trends in Cognitive Sciences*, *12*(11), 411–417. https://doi.org/10.1016/j.tics.2008.07.006
Sharples, S., Cobb, S., Moody, A., & Wilson, J. R. (2008). Virtual reality induced symptoms and effects (VRISE): Comparison of head mounted display (HMD), desktop and projection display systems. *Displays*, *29*(2), 58–69. https://doi.org/10.1016/J.DISPLA.2007.09.005
Slater, M., & Sanchez-Vives, M. V. (2016). Enhancing Our Lives with Immersive Virtual Reality. *Frontiers in Robotics and AI*, *3*, 1–47. https://doi.org/10.3389/frobt.2016.00074
Stepan, K., Zeiger, J., Hanchuk, S., Del Signore, A., Shrivastava, R., Govindaraj, S., & Iloreta, A. (2017). Immersive virtual reality as a teaching tool for neuroanatomy. *International Forum of Allergy and Rhinology*, *7*(10), 1006–1013. https://doi.org/10.1002/alr.21986
Stockwell, G. (2008). Investigating learner preparedness for and usage patterns of mobile learning. *ReCALL*, *20*(3), 253–270. https://doi.org/10.1017/S0958344008000232
Stockwell, G., & Hubbard, P. (2013). *Some emerging principles for mobile-assisted language learning*. Monterey, CA.
Thornton, P., & Houser, C. (2005). Using mobile phones in English education in Japan. *Journal of Computer Assisted Learning*, *21*(3), 217–228. https://doi.org/10.1111/j.1365-2729.2005.00129.x
Ulrich, F., Helms, N. H., Frandsen, U. P., & Rafn, A. V. (2019). Learning effectiveness of 360° video: experiences from a controlled experiment in healthcare education. *Interactive Learning Environments*, *4820*. https://doi.org/10.1080/10494820.2019.1579234
Usai, F., O Neil, K., & Newman, A. J. (2017). Design and empirical validation of effectiveness of LANGA, an online game-based platform for second language learning. *IEEE Transactions on Learning Technologies*, *1382*(c), 1–1. https://doi.org/10.1109/TLT.2017.2762688
Viberg, O., & Grönlund, Å. (2012). Mobile Assisted Language Learning: A literature review. *Proceedings of the 11th International Conference on Mobile and Contextual Learning*, 1–8.
Villani, D., Morganti, L., Carissoli, C., Gatti, E., Bonanomi, A., Cacciamani, S., … Riva, G. (2018). Students' acceptance of tablet PCs in Italian high schools: Profiles and differences. *British Journal of Educational Technology*, *49*(3), 533–544. https://doi.org/10.1111/bjet.12591
Vygotskiĭ, L. S., & Cole, M. (1978). *Mind in society : the development of higher psychological processes*. Cambridge:MA: Harvard University Press.
Wu, J., Guo, R., Wang, Z., Zeng, R., & Wu, J. (2019). Integrating spherical video-based virtual reality into elementary school students ' scientific inquiry instruction : effects on their problem-solving performance problem-solving performance, *4820*. https://doi.org/10.1080/10494820.2019.1587469