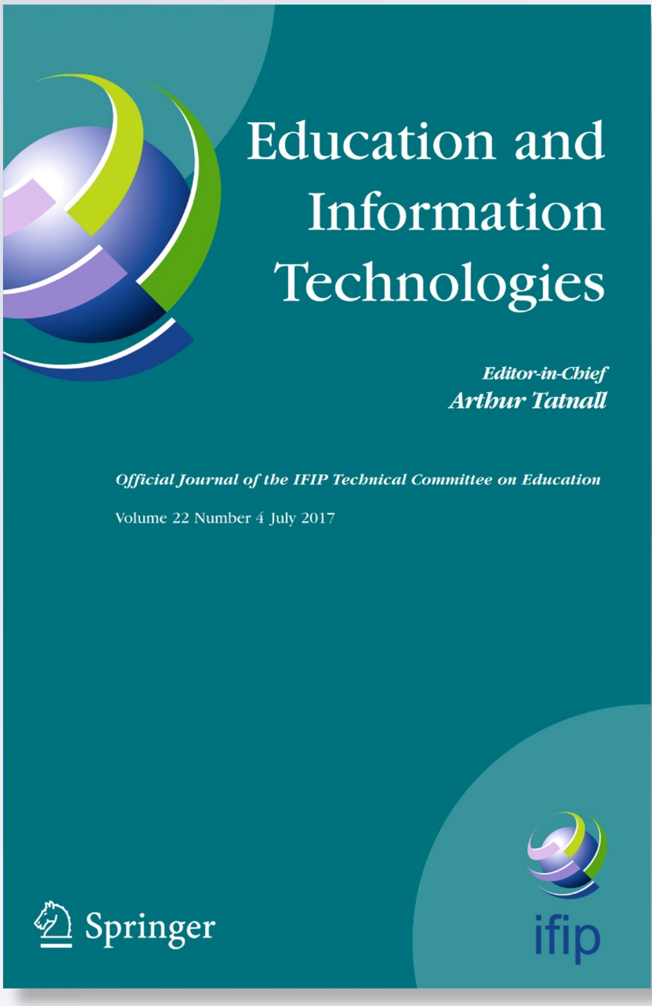
Content and language integrated learning in OpenSimulator project. Results of a pilot implementation in Greece

Emmanuel Fokides & Constantina Zampouli

Education and Information Technologies
The Official Journal of the IFIP Technical Committee on Education

ISSN 1360-2357
Volume 22
Number 4

Educ Inf Technol (2017) 22:1479-1496
DOI 10.1007/s10639-016-9503-z



Springer

Educ Inf Technol (2017) 22:1479–1496
DOI 10.1007/s10639-016-9503-z

Content and language integrated learning in OpenSimulator project. Results of a pilot implementation in Greece

Emmanuel Fokides[1] · **Constantina Zampouli**[1]

Published online: 20 May 2016

Abstract The study presents the results of a pilot implementation of the Content and Language Integrated Learning in OpenSimulator Project-CLILiOP. Content and Language Integrated Learning together with constructivism provided the theoretical basis of the project. A multi-user virtual environment was developed using Opensimulator, which entailed a tour of the students' avatars on a virtual island where they viewed and exchanged information in the English language regarding geographical terms and concepts. The project's duration was ten weeks. A total of 105 students participated, divided into three groups: one used the application, the second was taught the same cognitive material using a conventional teaching method, but still under the scope of CLIL, while the third was conventionally taught. Results indicate that the learning outcomes of CLILiOP were better compared to the other two teaching methods, but also underline the need to further investigate the uses of 3D multi-user virtual environments in second language learning.

Keywords CLIL · EFL · Opensimulator · 3D multi-user virtual environments · Constructivism · Virtual learning environments

1 Introduction

Content and Language Integrated Learning (CLIL), has become an important issue in the last few years in the English as a foreign/second language (EFL) debate. CLIL is a two-fold educational approach in which a language other than the learner's mother tongue is used for the learning and teaching of both content and language (Maljers

✉ Emmanuel Fokides
fokides@aegean.gr

[1] Department of Primary School Education, University of the Aegean, 1 Dimokratias str., 85132 Rhodes, Greece

2007). Withal, so as to procure this dual objective, the development of a special approach to teaching is needed so that the content is not simply taught in a foreign language but with and through a foreign language. This suggests a more integrated approach to both teaching and learning, entailing a much more thorough curriculum planning and implementation, not merely to the way languages are being taught, but to a holistic educational process in general (Eurydice Report 2006).

Contemporary educational practices attempt to take advantage of technologies which enable learners to easily seek and share information and promote self-motivated, autonomous and informal learning (Dabbagh Dabbagh and Kitsantas 2012). Creative expression, critical and reflective thinking, analysis and synthesis of information are also included in the digitalized arsenal of education (Trilling and Fadel 2009). The evolution of computer games and of 3D simulations have outlined even more novel means of communication, collaboration, and formation of social discourse, that can be exploited in education (Taiwo 2010). Technology has an impact on the learning process and, as a consequence, the learning outcomes are improved (Felix 2005). Studies conducted worldwide suggest quite a few positive effects on students' attitudes, skills, motivation, computer literacy and learning of a second language (Blake 2000; Liu et al. 2002). With the emergence of 3D multi-user virtual environments (MUVEs), a whole new world of learning-teaching is under way. Research has shown that MUVEs promote high levels of intrinsic motivation and creativity (Brown et al. 2010), due to the enjoyment and fun that the experiences provide (Chandra et al. 2009).

Although a sufficient amount of research has been contacted worldwide (e.g. Jauregi et al. 2011; Cenoz and Ruiz de Zarobe 2015), CLIL still constitutes a novelty in the Greek educational mainstream. Furthermore, and to the best of our knowledge, the combination of CLIL with a 3D virtual environment is yet to be explored. Towards this end, a project, named CLIL in Opensimulator Project (CLILiOP), was designed and implemented, upon which the particular research is based, as presented in the coming sections.

2 Content and language integrated learning-CLIL

CLIL was originally defined in 1994 (Marsh 1994) and launched in 1996 by UNICOM, University of Jyväskylä, Finland and the European Platform for Dutch Education. Coyle's (2008) "4Cs", namely, Content, Cognition, Communication and Culture/Citizenship are the framing principles to outline the pedagogic practice in CLIL (Fig. 1). The 4Cs framework suggests that effective CLIL is implemented through knowledge development, subject matter comprehension, engagement in cognitive processing, communicative discourse, and an intercultural awareness (Coyle 2008). Coyle's viewpoint underlines the fact that CLIL is about both appropriate and effective language usage. The main focus lies on the interrelationship between content (e.g. subject matter), communication (e.g. language), cognition (e.g. thinking) and culture (awareness of self and "otherness"), to build on the synergies of integrating learning (content and cognition) and language learning (communication and cultures) (Coyle 2008). At the heart of this cross-curricular approach lies the assertion that education should target -at once- the intellectual, emotional, and social child's development (Prentza 2013).

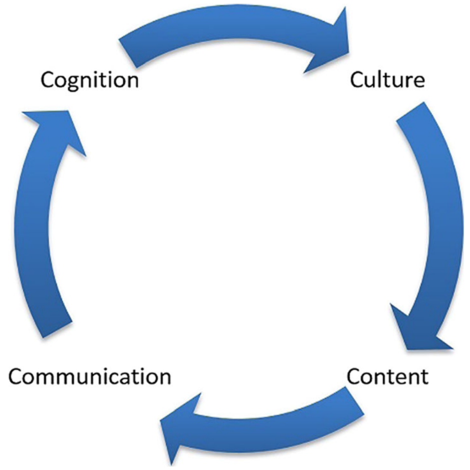


Fig. 1 The 4 C's framework

Marsh and Wolff (2007) put forward the view that there is no single model or an archetypical matrix that can be exported from one European country to another so as to implement CLIL. It seems to flourish in various settings, such as: (a) monolingual, where students, mostly non-native speakers, learn a subject through CLIL, (b) bilingual, where students learn 50 % or more curriculum subjects in a second or foreign language, (c) multilingual, where students learn curriculum subjects in three or more languages, and (d) plurilingual, where students learn several languages, one or more of which may be through CLIL (Marsh et al. 2010).

Students participating in CLIL groups significantly outstripped non-CLIL groups in pronunciation, vocabulary, grammar, fluency, and content (Lasagabaster 2008). The same holds true for text reconstruction, listening and reading comprehension, grammatical proficiency, writing, and socio-pragmatic competency (Klieme et al. 2006). In longitudinal studies, students in CLIL programs performed better on reading comprehension, general oral proficiency and on the final exams in English (Admiraal et al. 2006). Moreover, students taking two CLIL subjects had significantly better results than those taking just one (de Zarobe 2008). Evidence also suggests that there is a benefit of beginning at a younger age. If so, students demonstrated a wider range of specialized vocabulary, greater length of utterance, and spoke coherently at some length and with little hesitation (Dobson et al. 2010). On the other hand, CLIL is not officially implemented in Greek state schools. There may be sporadic attempts, mainly in state or private High and Junior High Schools, but only on an experimental level (Mpaltsavia 2011; Vlachos 2009).

3 3D multi-user virtual environments

As the Information Age progresses, new -digital- tools are integrated into the school environment, capturing students' imagination and leading to 21st-century skills: critical thinking, problem-solving, creativity, innovation, media literacy, ICT literacy, flexibility, initiative, and self-direction (Trilling and Fadel 2009). Virtual environments (VEs) first emerged in the '70s and later on, led to MUDs (Multi-user domains) and MOOs (Multi-user domains object-oriented), which language teachers were able to exploit for teaching foreign languages (Shield 2003). The turn of the century saw the emergence of 3D multi-user VEs (MUVEs). The most popular MUVE is Second Life (SL) which was launched in 2003 (http://secondlife.com/). In SL, the user can create his/her own avatar (virtual self) and explore unique virtual environments. SL soon grasped the educators' attention: social exchange, collaboration between avatars, visual and audio stimuli are but a few of its real-time advantages (Zheng and Newgarden 2011). In 2007, the OpenSimulator project (http://opensimulator.org/) was launched to take a further step: it is an open source MUVEs server, supporting several clients while maintaining compatibility with SL. The aforementioned applications are only but two out of a large number of similar or totally different technologies, all trying to exploit the potential of MUVEs.

Constructivism provides the theoretical framework for the educational uses of MUVEs (Dickey 2005; Kirkley and Kirkley 2005). Constructivism supports the notion that learners construct knowledge on the basis of what they already understand and as they make connections between new and old information (Ertmer and Newby 2013; Piaget 1985). Learners' prior ideas, experiences, and knowledge, interplay and may even clash with new experiences and their interpretations of the surrounding environment (Savery and Duffy 1995). Inconsistencies between what they already know and new persuasive information bring their current understanding into question and results in cognitive conflict (Bruner 1973). When students resolve these inconsistencies, they actually create new ways to reconcile their prior knowledge with the new information (Bruner 1973). Thus, the resolution of cognitive conflict drives learning.

By simulating real or imaginary environments, MUVEs give users the sense of immersion, of "being there" (Hew and Cheung 2010). MUVEs are used for constructivist learning because of the opportunities for learners to express their personal thoughts to explore, collaborate, to be immersed in the environment, and to construct their knowledge (Pan et al. 2006). They also attract the interest of students and, in combination with the sense of presence and the in-world activities, the educational process becomes more effective (Mikropoulos 2006; Martin et al. 2011).

4 Application's rationale and development

Given the aforementioned orientation, one can conclude that CLIL's approach in EFL is the interweaving of content and language, highlighting the construction of knowledge rather than the instruction itself. Also, MUVEs provide a potential new universe for students to develop language skills through media. Despite the fact that CLIL appears to play a core role in European educational trends, we noticed a relative lack of discussions in CLIL combined with MUVEs literature; therefore, we decided to focus our study exactly there.

We chose Geography as the subject matter of the project, and more specifically geographical landforms (i.e., peninsula, river delta, archipelago, etc.), world biomes (aquatic and terrestrial), longitude-latitude concepts (i.e., how latitude and longitude define locations on Earth, time zones, prime meridian, equator), and general geographical terms (i.e., tectonic plates, continents, and oceans). The specific topics were chosen since we considered easier for the EFL teacher to activate students' prior knowledge in their L1 (first language) and immerse them in the world of CLIL, even though it is sometimes difficult to transfer knowledge in a L2 or L3 (University of Cambridge 2014).

We considered several platforms for the development of the project's application, on the basis of the following criteria: (a) open source platform, (b) free client/server software, (c) easy content creation for non-experts, and (d) in-world web browsing and file sharing capabilities. We also considered additional features, such as instant messaging and voice chat (Table 1). From the comparison of these features, it became evident that Opensimulator was the best selection for the project at hand.

The application's total virtual space was 1024 × 1024 m. It included a main island with all the relevant landforms (e.g., a peninsula, a gulf, a gorge), biomes (e.g., a swamp, a forest, a lake, a river with a waterfall, a jungle), several landmarks (e.g., a castle, a pier, a beach), and a media corner, which served as the learning center for the project (Fig. 2). We had to balance realistic and detailed illustration of the virtual world in one hand, and not overloading it on the other. Overloading a 3D application quite easily leads to severe lagging when running on client computers. Once logged in, users would be able to explore the island and locate the signposts that would reveal further online information about a term or a concept (i.e. its definition, a synonym, examples and pronunciation), satisfying multiple types of intelligence, such as linguistic, musical, spatial, and interpersonal (Gardner 2011). Thus, pertaining to the compatibility of virtual environments to existing learning styles, the virtual environments provide multiple channels that satisfy students' learning styles (Junglas et al. 2007; Henderson et al. 2009). By doing so, we believe that there can be an increase in students' achievement, as Dunn et al. (2002) also pointed out.

Students were introduced to new vocabulary via the above-mentioned sign posts and the media corner which included videos as well as presentations. Video brings the outside world into the classroom adding a "reality" element to teaching and learning. Not only that but also video usage in the EFL classroom provides all the paralinguistic features of the language, namely body language, gestures, facial expressions, tone and pitch of voice (Potosí et al. 2009). Apart from the visual stimulus, research has shown that exposure to video activities improves students' listening skills, and it also

Table 1 Pros and cons of the different MUVEs platforms

	MUVEs platforms			
	Active worlds	Open cobalt	Second life	Opensimulator
Open source	no	yes	no	yes
Free client/server	yes/no	yes/yes	yes/no	yes/yes
Easy content creation	no	no	yes	yes
Web browsing	yes	yes	yes	yes
Voice chat	yes	yes	yes	yes
Instant messaging	yes	yes	yes	yes
Full functionality	yes	no	yes	yes

Fig. 2 Screenshots from the application

stimulates students' oral production practices (Potosí et al. 2009). In light of the above, we used videos as students' "shared experience". The students would be able to pause, rewind, or fast forward the video, take notes and actually use it as a tool for reference of language.

Despite criticism that presentations create passive students (Johnson and Sharp 2005), studies have argued that as long as students are not treated as *tabulae rasae*, presentations can be used as a tool to facilitate participation and discussion (Taylor n.d.). Using presentations for student-centered projects can be motivating and enjoyable not only for students but also for instructors (Apple and Kikuchi 2007). They foster students' sense of self-reliance, promote group work and scaffolding in second language learning (Taylor n.d.). Taking into consideration the above, the presentations embedded in the application were selected so as to: (a) introduce new vocabulary with thematic pictures, (b) present drill wh- questions (e.g. What…?, Where…?), (c) review previous vocabulary, and (d) elicit additional vocabulary on the topic with additional pictures.

The most rigorous stage -apart from the actual development of the virtual world per se- was the planning, designing, and collecting all necessary cognitive material and embedding it in the application. Ready-made scripts were used; therefore, scripting/programming was completed in a relatively short period of time. While one might argue that MUVEs may be time-consuming to produce, once created, teachers need only to change the content slightly so as to prepare future lesson plans. Table 2 illustrates the hours spent in designing and developing the application. It is worth mentioning that we initially tested the application with a small group of students in order to check its functionality and we made minor technical adjustments.

Table 2 The application's total development time (approximately)

Development stage	Hours
Cognitive material collection	40
Virtual world development	70
Images, videos, presentations	15
Scripts	20
Quality control	5
Minor adjustments	10
Total	160

As a final note, the application can be hosted on any school's computer which meets the hardware requirements so as to act as an Opensimulator server. By having the server running on school's local network, client computers do not face network lagging problems, the most common and persistent problem in all MUVEs, where data have to be routed to and from distant servers.

5 CLILiOP's design and implementation

The application, as described in the previous section, formed the basis of a project named "Content and Language Integrated Learning in Opensimulator Project, CLILiOP". CLILiOP consisted of three stages: a two-week pre-stage, a six-week while-stage, and a two-week post-stage. During the pre-stage, the participants received instructions pertaining to the project and explored the affordances and constraints of the application in order to proactively face difficulties while using it. The cognitive aims of the while-stage (weeks 3–8), covered the development of analytic and synthetic skills, learning strategies, visual and auditory perception. Pedagogical aims encompassed self and social growth in the form of enhancing ethos of communication, co-operation skills, and respect for oneself and others that may be linguistically or culturally different. The learning tasks for each week were one hour of learning activities in Opensimulator, followed by one-hour in-classroom learning activities. The learning activities in Opensimulator included virtual tours in groups, small group discussions, and peer-group learning. In-classroom activities gave students the chance to practice the vocabulary and language forms by discussing, role playing, writing reports, and completing gap-fill exercises. The post-stage lasted two weeks (weeks 9–10) and included a group-reflection with discussion and note-exchanging, as well as a self-reflection report. It culminated in "Opensimulator Evaluation", where each student evaluated different aspects of the project. Metacognition was the key element in the post-stage tasks. CLILiOP's stages and the relevant tasks are presented in Table 3.

In parallel with CLILiOP, a second project was developed, which followed the same structure as CLILiOP, but without the pre-stage. Wherein, we followed a more conventional teaching approach but still under the scope of CLIL. The same presentations, videos, cognitive and pedagogical aims, and the same learning tasks were used, excluding the virtual world element. The two-hour learning

Table 3 CLILiOP's stages and tasks

	In-world tasks/activities	In-classroom tasks/activities
Week 1 (pre-stage)	Login and logout of Opensimulator. Get accustomed with navigation. Group split-up (group-work & individually)	Preparatory workshop: "What is Opensimulator?" (orally, group-work with researcher).
Week 2 (pre-stage)	Get accustomed with navigation. Follow/give instructions (group-work & individually).	
Week 3 Landmarks	Take a tour. Pinpoint on virtual map key-locations/landmarks. Visit key locations. Group discussion (instant messaging-IM).	Group discussion. Gap-fill exercises, write a report about your virtual tour (individually).
Week 4 Geography terms	Virtual tour. Locate media corner. Watch presentation. Locate sign posts, get more info. Group discussion (IM).	Group discussion. Write a report on geography terms (individually).
Week 5 Biomes I	Virtual tour. Locate and examine biomes. Watch video. Note-taking. Watch presentation. Group discussion (IM).	Group discussion, role playing (orally). Gap-fill exercises, write a report on biomes (individually).
Week 6 Biomes II	Virtual tour. Locate presentation screen. Watch the presentation on biomes. Locate and examine biomes. Note-taking. Group discussion (IM).	Group debate: "Which is my favorite biome and why?" (orally).
Week 7 Longitude and latitude	Virtual tour. Watch video. Note-taking. Locate sign posts, get more info. Guide team to landmarks using the virtual map. Group discussion (IM).	Group discussion. Use provided worksheet and give instructions using longitude and latitude terms (orally). Write a report on longitude and latitude (individually).
Week 8 Landforms	Virtual tour. Locate/visit the different landforms. Locate sign posts, get more info. Group discussion (IM).	Group discussion. Write a report on the landforms you visited (individually).
Week 9 (post stage)	Final group meeting. Revisit key-locations. Group-reflection: meet and discuss your experience (IM).	Group-reflection: meet exchange notes (written-orally). Write report: "What have I learned in the virtual world?" (individually).
Week 10 (post stage)		Opensimulator evaluation (individually).

Two-hour session per week

activities for each week were video and presentations watching, small group discussions, peer-group learning, and lectures. Table 4 illustrates the conventional CLIL project's stages and tasks.

The pilot implementation of CLILiOP lasted for ten weeks; from September 21 to November 30, 2015, and the participants were 35 students in the 6th (last) grade of a primary school in Athens. The EFL teacher was on site to provide learning or technical support (Fig. 3). Since the computer lab had limited number of computers, CLILiOP's students were split into 3 groups, so that each student used his/

Table 4 Conventional CLIL's stages

	Tasks/Activities
Week 1 Landmarks	Use provided worksheet on landmarks. Pinpoint on map key-locations. Group discussion (orally). Gap-fill exercises, write a report about your country (individually).
Week 2 Geography terms	Use provided worksheet on geography terms. Watch presentation. Group discussion. Write a report on geography terms (individually).
Week 3 Biomes I	Use provided worksheet on biomes. Watch video. Note-taking. Watch presentation. Group discussion, role playing (orally). Gap-fill exercises, write a report on biomes (individually).
Week 4 Biomes II	Use provided worksheet on biomes. Group debate: "Which is my favorite biome and why?" (orally).
Week 5 Longitude and latitude	Use provided worksheet on longitude and latitude. Watch video. Note-taking. Group discussion. Use provided worksheet and give instructions using longitude and latitude terms (orally). Write a report on longitude and latitude (individually).
Week 6 Landforms	Use provided worksheet on landforms. Locate the different landforms on the map. Group discussion. Write a report on the landforms (individually).
Week 7 (post stage)	Group-reflection: meet exchange notes (written-orally). Write report: "What have I learned so far?" (individually).
Week 8 (post stage)	Conventional CLIL's Evaluation (individually).

Two-hour session per week

her own computer. As a result, each session was repeated 3 times. At the same period of time, conventional CLIL was also implemented. Once again, participants were 35, 6th-grade students of a nearby school.

CLILiOP almost immediately draw the attention of students. They were very interested in exploring, locating all the sign posts, and in finding all the details of the virtual environment. By doing so, they were also highly motivated in

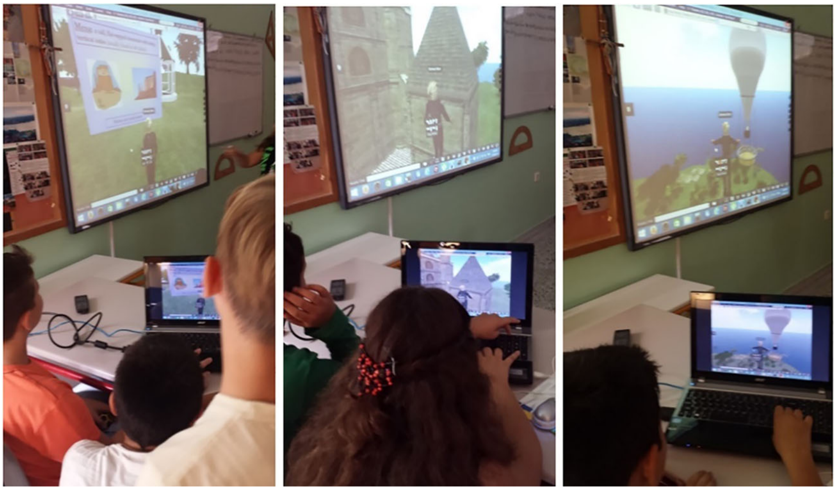


Fig. 3 Screenshots from CLILiOP's implementation

communicating and collaborating in order to achieve their goals. Nevertheless, code-switching from English to Greek and back was inevitable. For instance, a group had the following exchange in instant messaging:

Student A: <*Go left. Now turn right…*>

Student B: <*Where is the tambela?*>*

Student A: <*Look! Over there.*>

*tambela is the Greek word for sign post

6 Research design and procedure

Students in CLILiOP and in the conventional CLIL were the study's initial target groups. On the other hand, there were two facts that we had to take into consideration when designing the research procedure. EFL is officially taught in the Greek educational system from the 1st grade. Pupils initially develop their listening and speaking skills, whereas from the 3rd grade onwards they also develop their reading and writing skills. As already mentioned, 6th grade students participated to the study. This means that the participants already had a five-year experience of English. In addition, many students study English in private evening schools or are home tutored. Therefore, participants' knowledge level of English may considerably vary. We were also interested in testing the efficiency of both conventional CLIL and CLILiOP. To account for the above, we decided to include a control group. No special provisions were taken for this group; there were no pre- and post-stages, students were taught according to the official guidelines, by using the course book and more specifically the unit that contained the exact same cognitive material as in CLILiOP. The only deviation from the official guidelines was the addition of more cognitive material and exercises/tasks, so as to keep step with the other groups. Thirty-five 6th grade students of a school from the same area as the other two formed this group. Therefore, a total of 105 students participated in the study (35 in each group).

For data collection purposes, we devised a total of 5 questionnaires and tests:

- A questionnaire attempting to gather data regarding the participants' linguistic profile, as well as demographic data (9 items-all groups).
- The above questionnaire was administered together with a pre-test with lexical (vocabulary and grammar) questions (22 items-all groups). The purpose of this test was to check whether all groups had the same initial knowledge level of English.
- A post-test with lexical (vocabulary and grammar) questions (22 items-all groups). The purpose of this test was to gain insight regarding the effectiveness of all teaching methods.
- A questionnaire attempting to gather data regarding students' views and experiences with CLILiOP and the related activities (20 items-only CLILiOP group).

- A follow-up test was carried out 4 weeks after the projects' completion (all groups). The purpose of this test was to gain insight regarding the metacognitive results of all teaching methods. There were no questions in this test; students were asked to use as many terms presented during lessons as they could remember, structuring their own sentences or phrases.

Both pre- and post-tests used Yes-No, open-ended and multiple-choice questions. The second questionnaire had open-ended as well as Likert-type questions; statements were set on a five-point scale ranging from "strongly disagree" to "strongly agree".

The aforementioned procedure was to examine the following hypotheses:

H1: CLILiOP produces significantly better cognitive results compared to a conventional CLIL and a conventional teaching method.
H2: CLILiOP produces significantly better metacognitive results compared to a conventional CLIL and a conventional teaching method.

7 Results analyses

As already mentioned in the previous section, a total of 105 6th grade students, in 3 neighboring schools participated in the study, divided into 3 groups of 35 (control/conventional, conventional CLIL, and CLILiOP). There were 16 boys and 19 girls in both the control and the CLILiOP groups while the conventional CLIL group had 18 boys and 17 girls. Almost all (except 7), were learning English either in private evening schools or were home tutored, in parallel with learning English at school. In addition, their proficiency abilities ranged from junior/basic users (A1-A2) to intermediate users (B1) according to the Common European Framework of Reference for Languages (Council of Europe 2011). For the analysis of the results in pre- and post-tests, scores were computed, on the basis of the number of correct answers in the relevant questionnaires ($min = 0$, $max = 22$, for both tests). Mean scores per group of participants, and per test, are presented in Table 5.

A One-way ANOVA was conducted to compare the total scores of the three groups in the pre-test, in order to check if these groups had any significant differences. The same procedure was followed to compare the effects of the type of teaching method on the total scores in post-test. Prior to conducting these tests, we checked whether the

Table 5 Means and standard deviations on pre and post-test

Pre-test						Post-test					
Control group ($N = 35$)		Conventional CLIL group ($N = 35$)		CLILiOP group ($N = 35$)		Control group ($N = 35$)		Conventional CLIL group ($N = 35$)		CLILiOP group ($N = 35$)	
M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
7.14	2.30	6.57	2.18	7.23	2.42	6.95	4.34	12.14	3.27	17.38	4.56

assumptions of ANOVA testing are violated: (a) all groups had the same number of participants ($N = 35$), (b) there were no outliers and the data was normally distributed, as assessed by Q-Q plots and Shapiro-Wilk test ($p > .05$ in all cases), and (c) homogeneity of variance was also not violated, as assessed by Levene's test ($p = .60$ for pre-test and $p = .15$ for post-test). Since all assumptions were met, the analysis could be conducted. An analysis of variance showed that there were no significant differences on total scores in the pre-test, $F(2, 102) = 0.51$, $p = .60$. On the other hand, the type teaching method had a significant effect on total scores in post-test, $F(2, 102) = 33.97$, $p < .001$.

Post-hoc comparisons using the Tuckey HSD test were conducted on all possible pairwise contrasts in post-test. All pairs of groups were found to be significantly different at the $p < .05$ level. The mean total score in post-test for the CLILiOP group ($M = 17.38$, $SD = 4.56$) was significantly higher than that of the conventional CLIL group ($M = 12.14$, $SD = 3.27$) and both were significantly higher than that of the control group ($M = 6.95$, $SD = 4.34$).

Taken together, these results suggest that:

- All groups had the same knowledge level of the English language prior to conducting the research since they did not have statistically significant differences in the pre-test. As all groups had the same initial starting point, any differences observed in the participants' knowledge acquisition after the interventions, can be attributed to the different teaching methods that were followed.
- The type of teaching method used had a statistically significant effect on total scores in post-test. More specifically, students participating in CLILiOP had statistically significant better cognitive results compared to both the conventional teaching and to conventional CLIL, with all having the same cognitive content. Therefore, H1 is confirmed.

Two scores were computed on the follow-up test: (a) the total number of terms used, and (b) the total number of wrong uses and of spelling mistakes in these terms. The initial screening of data revealed that the scores of the control group were very low, meaning that students used very few terms ($M = 9.2$, $SD = 5.95$) and made quite a lot of mistakes ($M = 7.15$, $SD = 1.38$). Since it was more than clear that the control group fell short compared to the other two groups, it was not included in the analysis. Mean scores per group of participants, and per test, are presented in Table 6.

Table 6 Means and standard deviations on follow-up test

Terms used				Mistakes			
Conventional CLIL group ($n = 35$)		CLILiOP group ($n = 35$)		Conventional CLIL group ($n = 35$)		CLILiOP group ($n = 35$)	
M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
16.1	5.53	21.28	6.57	2.38	0.97	2.67	1.56

Prior to conducting statistical analysis on the results, we checked whether the assumptions of ANOVA testing are violated. We found that: (a) all groups had the same number of participants ($n=35$), (b) there were no outliers, (c) data was normally distributed in the number of terms used, as assessed by Q-Q plots and Shapiro-Wilk test ($p=.98$), (d) homogeneity of variance in the number of terms used was also not violated, as assessed by Levene's test ($p=.28$), (e) there were problems regarding the normality of data in mistakes, as assessed by Q-Q plots and Shapiro-Wilk test ($p=.007$), and (f) homogeneity of variance was also violated in mistakes, as assessed by Levene's test ($p=.014$).

Since all assumptions were met for ANOVA testing on the scores regarding the number of terms used, the analysis could be conducted. An analysis of variance showed that the teaching method, had a significant effect on the total number of terms used, $F(1, 68)=7.66$, $p=.009$.

Like other parametric tests, the analysis of variance assumes that the data fit the normal distribution. On the other hand, the literature suggests that ANOVA is not very sensitive to moderate deviations from normality and the false positive rate is not affected very much by this violation (Glass et al. 1972; Lix et al. 1996). However, since it is debatable under which conditions and with which techniques it is acceptable to conduct an ANOVA in non-normally distributed data, we decided to analyze the scores of mistakes using the Kruskal-Wallis H test, which is a non-parametric test. Even though this test does not assume that the data fit the normal distribution, it assumes that the data in different groups have similarly shaped distributions (Corder and Foreman 2009; Siegel and Castellan 1988), as was the case in scores in mistakes. The test revealed that the distribution of mistakes was the same across groups and not statistically significantly different ($p=.70$). One way to interpret this finding is the small number of mistakes made in both groups. Combined with the results regarding the number of terms used, one can support the notion that although the same number of mistakes was made in both groups, the CLILiOP group used significantly more geographical terms. Therefore, H2 can also be accepted.

All but a few students (5 cases) use computers at home (mainly to play games), so it is safe to assume that the application was compatible with their computer skills. Students made positive remarks regarding their experiences while using CLILiOP's application. In particular, all of them agreed that it was attention-grabbing, fun to use/game-like and that they would like to have more time to explore the virtual world. They also enjoyed the in-world (31 cases), as well as the in-classroom (28 cases) activities. Moreover, they stated that the application helped them in having a better understanding of subjects related to geography (29 cases). Most students (32 cases) liked the communication aspect of the application; the chance to express themselves and to talk to each other through chat. A student distinctly noted that the application gave her the opportunity to talk and write freely in English as opposed to her reservations when being taught with "the other way", meaning conventionally. In contrast, another student expressed her hesitation writing on chat because she might make many mistakes and other students might notice this. Similar fears were expressed by other learners (4 cases) who, however, highlighted that practice would help cope with this restraint.

On the negative side, a number of CLILiOP's participants stated they encountered some kind of technical and/or usage problems. These were: unsatisfactory application's display speed due to hardware issues-outdated computer (11 cases), orientation problems (4 cases), and handling the avatar (3 cases). Students that faced orientation problems and problems using the application were the ones that do not use computers at home or use them but do not play games. Nevertheless, these problems were noticed by the EFL teacher, who offered help and the problems were overcome after the third session.

8 Discussion

Our main concern was the initial knowledge level of the English language in all participating groups, given that the vast majority of students are home tutored or study in private evening schools, in parallel with learning English at school. The pre-test confirmed that all groups had the same initial starting point since they did not have any statistically significant differences. The post-test revealed that both CLIL teaching methods (conventional CLIL and CLILiOP) produced very good results. Conventional CLIL group had almost twice the scores of conventional teaching ($M_{CLIL} = 12.14$, $M_{conventional} = 6.95$), while CLILiOP group had more than twice ($M_{CLILiOP} = 17.38$). Therefore, it was a matter of comparing the two CLIL methods. This comparison shown that CLILiOP surpassed conventional CLIL in two ways: (a) scores in post-test were statistically significantly different in favor of CLILiOP, and (b) although students in both groups made more or less the same number of mistakes in the follow-up test, CLILiOP's students used statistically significant more geographical terms ($M_{CLIL} = 16.1$, $M_{CLILiOP} = 21.28$).

As a result of the above, both research hypotheses were confirmed; CLILiOP produced significantly better cognitive and metacognitive results compared to a conventional CLIL and a conventional teaching method. It is important to highlight the significance but also the constraints of these findings:

- We need to find better ways to teach EFL. Textbooks are somehow obsolete and conventional teaching proved to be ineffective; mean scores in both pre- and post-tests were almost identical ($M_{pre} = 7.14$, $M_{post} = 6.95$) and scores on post-test were, by far, worse than the other two methods. In addition, CLILiOP surpassed the conventional CLIL which was also based on printed notes and reading material. These findings are in line with other studies pointing to the same direction (e.g. Lasagabaster 2008; Klieme et al. 2006; Admiraal et al. 2006; Alonso et al. 2008).
- When it comes to choosing an alternative teaching method, it appears that CLIL is the desired one. It seems to be a method that can be applied in diverse educational contexts and views students' needs as a whole (Coyle 2008; Prentza 2013). All these aspects of CLIL were utilized when designing both conventional CLIL and CLILiOP. Results in both, are in favor of the above assertion.
- CLILiOP surpassed conventional CLIL, but due to the small sample size, only assumptions can be made as to the exact reason this happened. Certainly the application was compatible with most students' computers skills. Even though a few students faced some problems, they were the ones who were not allowed to use computers at home, thus, were not easily adapted to the virtual world. Nevertheless, after the first three sessions, all students were able to navigate their avatars in the virtual environment with ease.

- Students enjoyed the game-like characteristics of the application, which was one of the reasons for attracting their attention, as others also note (Canfield 2008; Cooke-Plagwitz 2008; Chan 2008). This might lead to increased motivation since enjoyment and fun are experiences that are provided consistently in MUVEs (Chandra et al. 2009).
- Students also stated that they would like to explore more the virtual environment, but also that they felt more free to communicate with each other. The immersive nature of the virtual world encouraged users to interact with each other, and this, in turn, stimulated communication and social practices (Panteli 2009).
- The fact that CLILiOP students outperformed the conventional CLIL students in the follow-up test, indicates that learning in MUVEs has, at least, to some extent, long-lasting results and also an effect on students' metacognitive knowledge and skills as Nelson and Ketelhut (2007) also suggested.
- Having taken into consideration the above, our view tends to be that the difference between conventional CLIL and CLILiOP cannot be attributed to just one or more MUVEs characteristics, but to all of them and to the way that CLILiOP was organized.
- Geography is suitable for 3D visualization, but not all school subjects are. The starting point is also important. A MUVEs application might be pleonastic to students with a good knowledge level of a given subject.

The application's good results come at a price; development is indeed a time-consuming process. Then again, the extra time allocated to the development is rewarded by time gained in class; there is extra free time to monitor students' progress and provide feedback, leading to more attentive, on-task and focused students, therefore possibly better learning results. The biggest difficulties encountered during the implementation of the project were technical issues and a few usage problems. The school's computer lab had some outdated computers, which only meant crashing downs and laggings, making the students restless. Technical problems might lead to the obstruction of the learning process and users' loss of interest (Coban et al. 2015). This issue is common in all 3D applications; they require quite powerful computers to display complex graphics. Even though we tried to balance detailed illustration of the virtual environment and performance, outdated computers are *de facto* problematic and not much can be done.

9 Conclusion

The current study emerged from the need to investigate how MUVEs could be used in combination with CLIL in EFL teaching. Towards this end, an application was developed using Opensimulator so that a group of students, participating in the ten-week project CLILiOP, could virtually tour and study geography related concepts. The same cognitive material was given to another group of students, but this time, a more conventional CLIL teaching method was used. In order to compare results, we included a third group of students, which simply studied the textbook and was conventionally taught. Data analyses revealed that students participating in CLILiOP had better cognitive and metacognitive results. What is more, and according to students' comments, they were actively engaged and enjoyed using the application. As a result, it can be argued that the combination of CLIL and MUVEs, constitutes an effective educational tool, which provided an authentic contextualized game-like learning environment, where learners could express themselves, communicate with each other and interact.

On the other hand, the study has limitations that need to be acknowledged. Even though all necessary precautions were taken, we cannot be certain whether the tests and questionnaires accurately recorded students' knowledge and views. Since it was a pilot study, it was limited to a small number of participating students, therefore, its results cannot be generalized. Further studies are needed in order to identify differences or similarities to the findings of the present study. Research can also be conducted with a different timetable, age, and level group, to investigate possible advantages or disadvantages in teaching/learning.

Taking all limitations into consideration and in conclusion, it is our belief that CLIL together with MUVEs has a highly promising potential. It should be noted that CLILiOP is a work in progress. The goal is to further investigate different scenarios, under different situations and settings. However, the experimental data that were obtained, reinforce our view that MUVEs have a positive impact on EFL learning/teaching.

References

Admiraal, W., Westhoff, G., & de Bot, K. (2006). Evaluation of bilingual secondary education in the Netherlands: Students' language proficiency in english 1. *Educational Research and Evaluation, 12*(1), 75–93. doi:10.1080/13803610500392160.

Alonso, E., Grisaleña, J., & Campo, A. (2008). Plurilingual education in secondary schools: Analysis of results. *In Depth, 1*, 4.

Apple, M., & Kikuchi, K. (2007). Practical PowerPoint group projects for the EFL classroom. *The JALT CALL Journal, 3*(3), 110–122.

Blake, R. (2000). Computer mediated communication: A window on L2 Spanish interlanguage. *Language Learning & Technology, 4*(1), 120–136.

Brown, E., Hobbs, M., & Gordon, M. (2010). A virtual world environment for group work. *Novel Developments in Web-Based Learning Technologies: Tools for Modern Teaching: Tools for Modern Teaching, 233*.

Bruner, J. S. (1973). *Beyond the information given: Studies in the psychology of knowing*. New York: WW Norton.

Canfield, D. (2008). Using Immersive Learning Environments in Foreign Language Classes: Second Life. *CALICO Journal, 26*(1).

Cenoz, J., & Ruiz de Zarobe, Y. (2015). Learning through a second or additional language: Content-based instruction and CLIL in the twenty-first century. *Language, Culture and Curriculum, 28*(1), 1–7. doi:10.1080/07908318.2014.1000921.

Chan, J. (2008). Developing a Meaning-focused and Task-based Virtual Learning Reality. *CALICO Journal, 26*(1).

Chandra, S., Theng, Y. L., & Shou-Boon, S. F. (2009). Proposed theoretical framework for virtual world adoption. *SLACTIONS, 2009*, 22.

Coban, M., Karakus, T., Karaman, A., Gunay, F., & Goktas, Y. (2015). Technical problems experienced in the transformation of virtual worlds into an education environment and coping strategies. *Journal of Educational Technology & Society, 18*(1), 37–49.

Cooke-Plagwitz, J. (2008). Conversing in the Metaverse: Language Teaching and Learning in Second Life. *CALICO Journal, 26*(1).

Corder, G. W., & Foreman, D. I. (2009). *Nonparametric statistics for non-statisticians: a step-by-step approach*. John Wiley & Sons. doi: 10.1002/9781118165881.
Council of Europe (2011). *Common European Framework of Reference for Languages: Learning, Teaching, Assessment*. Council of Europe.
Coyle, D. (2008). CLIL-A pedagogical approach from the European perspective. In *Encyclopedia of language and education (pp. 1200–1214)*. Springer US. doi: 10.1007/978-0-387-30424-3_92.
Dabbagh, N., & Kitsantas, A. (2012). Personal learning environments, social media, and self-regulated learning: A natural formula for connecting formal and informal learning. *The Internet and Higher Education, 15*(1), 3–8. doi:10.1016/j.iheduc.2011.06.002.
de Zarobe, Y. R. (2008). CLIL and foreign language learning: A longitudinal study in the Basque country. *In Depth, 3*(4), 5.
Dickey, M. D. (2005). Brave new (interactive) worlds: A review of the design affordances and constraints of two 3D virtual worlds as interactive learning environments. *Interactive Learning Environments, 13*(1–2), 121–137. doi:10.1080/10494820500173714.
Dobson, A., Pérez, M., & Johnstone, R. (2010). *Bilingual Education Project (Spain): Evaluation Report*.
Dunn, R., Beaudry, J. S., & Klavas, A. (2002). What we know about how people learn. *California Journal of Science Education, 2*(2), 75–79.
Ertmer, P. A., & Newby, T. J. (2013). Behaviorism, cognitivism, constructivism: Comparing critical features from an instructional design perspective. *Performance Improvement Quarterly, 26*(2), 43–71. doi:10.1002/piq.21143.
Eurydice. (2006). *Content and language integrated learning (CLIL) at school in Europe*. Brussels: Eurydice European Unit.
Felix, U. (2005). Analysing recent CALL effectiveness research: Towards a common agenda. *Computer Assisted Language Learning, 18*(1), 1–32.
Gardner, H. (2011). *Frames of mind: The theory of multiple intelligences*. New York: Basic books.
Glass, G. V., Peckham, P. D., & Sanders, J. R. (1972). Consequences of failure to meet assumptions underlying fixed effects analyses of variance and covariance. *Review of Educational Research, 42*, 237–288. doi:10.3102/00346543042003237.
Henderson, M., Huang, H., Grant, S., & Henderson, L. (2009). Language acquisition in second life: Improving self-efficacy beliefs. In R. J. Atkinson, & C. McBeath (Eds.), *Same places, different spaces. Proceedings ascilite Auckland 2009* (pp. 464–474). Auckland: ascilite.
Hew, K. F., & Cheung, W. S. (2010). Use of three-dimensional (3-D) immersive virtual worlds in K-12 and higher education settings: A review of the research. *British Journal of Educational Technology, 41*(1), 33–55. doi:10.1111/j.1467-8535.2008.00900.x.
Jauregi, K., Canto, S., de Graaff, R., Koenraad, T., & Moonen, M. (2011). Verbal interaction in second life: Towards a pedagogic framework for task design. *Computer Assisted Language Learning, 24*(1), 77–101. doi:10.1080/09588221.2010.538699.
Johnson, K., & Sharp, V. (2005). Is PowerPoint crippling our students? *Learning & Leading with Technology, 33*(3), 6–7.
Junglas, I. A., Johnson, N. A., Steel, D. J., Abraham, D. C., & Loughlin, P. M. (2007). Identity formation, learning styles and trust in virtual worlds. *ACM SIGMIS Database, 38*(4), 90–96. doi:10.1145/1314234.1314251.
Kirkley, S. E., & Kirkley, J. R. (2005). Creating next generation blended learning environments using mixed reality, video games and simulations. *TechTrends, 49*(3), 42–53. doi:10.1007/BF02763646.
Klieme, E., Eichler, W., Helmke, A., Lehmann, R. H., Nold, G., Rolff, H. G., … & Willenberg, H. (2006). Unterricht und kompetenzerwerb in Deutsch und Englisch. Zentrale Befunde der Studie Deutsch-Englisch-Schülerleistungen-International (DESI) [Education and skills in German and English. Key findings of the study German English Student Services International]. *Eine Studie im Auftrag der Kultusminister der Länder in der Bundesrepublik Deutschland*.
Lasagabaster, D. (2008). Foreign language competence in content and language integrated courses. *The Open Applied Linguistics Journal, 1*(1), 30–41. doi:10.2174/1874913500801010030.
Liu, M., Moore, Z., Graham, L., & Lee, S. (2002). A look at the research on computer-based technology use in second language learning: A review of the literature from 1990–2000. *Journal of Research on Technology in Education, 34*(3), 250–273. doi:10.1080/15391523.2002.10782348.
Lix, L. M., Keselman, J. C., & Keselman, H. J. (1996). Consequences of assumption violations revisited: A quantitative review of alternatives to the one-way analysis of variance F test. *Review of Educational Research, 66*, 579–619.
Maljers, A. (Ed.). (2007). *Windows on CLIL: Content and language integrated learning in the European spotlight*. European Platform for Dutch Education.

Marsh, D. (1994). Bilingual education and content and language integrated learning. Paris, France: International Association for Cross-cultural Communication. *Language Teaching in the Member States of the European Union (Lingua), University of Sorbonne*.

Marsh, D., & Wolff, D. (2007). *Diverse contexts-converging goals. CLIL in Europe*. Frankfurt: Peter Lang.

Marsh, D., Mehisto, P., Wolff, D., & Frigols Martín, M. J. (2010). *European framework for CLIL teacher education: A framework for the professional development of CLIL teachers*.

Martin, S., Diaz, G., Sancristobal, E., Gil, R., Castro, M., & Peire, J. (2011). New technology trends in education: Seven years of forecasts and convergence. *Computers & Education, 57*(3), 1893–1906. doi:10.1016/j.compedu.2011.04.003.

Mikropoulos, T. A. (2006). Presence: A unique characteristic in educational virtual environments. *Virtual Reality, 10*(3), 197–206. doi:10.1007/s10055-006-0039-1.

Mpaltsavia, M. M. (2011). *Content and Language Integrated Learning in a Greek State Junior High School Class* (Master's Thesis, Hellenic Open University).

Nelson, B. C., & Ketelhut, D. J. (2007). Scientific inquiry in educational multi-user virtual environments. *Educational Psychology Review, 19*(3), 265–283. doi:10.1007/s10648-007-9048-1.

Pan, Z., Cheok, A. D., Yang, H., Zhu, J., & Shi, J. (2006). Virtual reality and mixed reality for virtual learning environments. *Computers & Education, 30*(1), 20–28. doi:10.1016/j.cag.2005.10.004.

Panteli, N. (2009). Virtual social networks: A new dimension for virtuality research. *Virtual Social Networks: Mediated, Massive and Multiplayer Sites*, 1–17. doi: 10.1057/9780230250888.

Piaget, J. (1985). *The equilibration of cognitive structures: The central problem of intellectual development*. Chicago: University of Chicago Press.

Potosí, A., Jhoana, L., Guarín Loaiza, E. A., & López García, A. C. (2009). *Using video materials as a teaching strategy for listening comprehension.*

Prentza, A. (2013). From teaching English to adults to teaching children: the changes that a teacher has to make. *MENON*, 57.

Savery, J. R., & Duffy, T. M. (1995). Problem based learning: An instructional model and its constructivist framework. *Educational Technology, 35*(5), 31–38.

Shield, L. (2003). MOO as a language learning tool. *Language learning online: Towards best practice*, 97–122.

Siegel, S., & Castellan, N. J., Jr. (1988). *Nonparametric statistics for the behavioral sciences*. New York: McGraw-Hill.

Taiwo, R. (Ed.). (2010). *Handbook of Research on Discourse Behavior and Digital Communication: Language Structures and Social Interaction: Language Structures and Social Interaction*. IGI Global. doi: 10.4018/978-1-61520-773-2.

Taylor, G. (n.d.) *Making a Place for PowerPoint in EFL Classrooms*. Retrieved from http://jaltcue.org/files/OnCUE/OCJ6-1/OCJ61%20pp%2041-51%20Taylor.pdf.

Trilling, B., & Fadel, C. (2009). *21st century skills: Learning for life in our times*. San Francisco: John Wiley & Sons.

University of Cambridge-ESOL (2014). *Teaching Geography through English-a CLIL approach*. Retrieved from: http://www.unifg.it/sites/default/files/allegatiparagrafo/21-01-2014/teaching_geography_through_clil.pdf.

Vlachos, K. (2009). The potential of information communication technologies (ICT) in content and language integrated learning (CLIL): the case of English as a second/foreign language. *CLIL practice: perspectives from the field. Juvaskyla: University of Jyvaskyla*, 189–198.

Zheng, D., & Newgarden, K. (2011). *Rethinking language learning: Virtual Worlds as a Catalyst for Change*.